

**INTERNATIONAL JOURNAL OF MODERN
ENGINEERING RESEARCH (IJMER)**

ISSN : 2249-6645



Volume 2 - Issue 5

Web : www.ijmer.com

Email : ijmer.editor@gmail.com

International Journal of Modern Engineering Research (IJMER)

Editorial Board

Executive Managing Editor

Prof. Shiv Kumar Sharma
India

Editorial Board Member

Dr. Jerry Van
Department of Mechanical, USA

Dr. George Dyrud
Research centre dy. Director of Civil Engineering, New Zealand

Dr. Masoud Esfal
R& D of Chemical Engineering, Australia

Dr. Nouby Mahdy Ghazaly
Minia University, Egypt

Dr. Stanley John
Department of Textile Engineering, United Kingdom

Dr. Valfitaf Rasoul
Professor and HOD of Electromechanical, Russian

Dr. Mohammed Ali Hussain
HOD, Sri Sai Madhavi Institute of Science & Technology, India

Dr. Manko dora
Associate professor of Computer Engineering, Poland

Dr. Ahmed Nabih Zaki Rashed
Menoufia University, Egypt

Ms. Amani Tahat
Ph.D physics Technical University of Catalonia-Spain

Associate Editor Member

Dr. Mohd Nazri Ismail
University of Kuala Lumpur (UniKL), Malaysia

Dr. Kamaljit I. Lakhtaria
Sir Padmapat Singhaniya University, Udaipur

Dr. Rajesh Shrivastava
Prof. & Head Mathematics & computer Deptt. Govt. Science & commerce College Benazir. M.P

Dr. Asoke Nath
Executive Director, St. Xavier's College, West Bengal, India

Prof. T. Venkat Narayana Rao
Head, CSE, HITAM Hyderabad

Dr. N. Balasubramanian
Ph. D (Chemical Engg), IIT Madras

Jasvinder Singh Sadana
M. TECH, USIT/GGSIPU, India

Dr. Bharat Raj Singh

Associate Director, SMS Institute of Technology, Lucknow

DR. RAVINDER RATHEE

C. R. P, Rohtak, Haryana

Dr. S. Rajendran

Research Supervisor, Corrosion Research Centre Department of Chemistry, GTN Arts College, Dindigul

Mohd Abdul Ahad

Department of Computer Science, Faculty of Management and Information Technology, Jamia Hamdad, New Delhi

Kunjai Mankad

Institute of Science & Technology for Advanced Studies & Research (ISTAR)

NILANJAN DEY

JIS College of Engineering, Kalyani, West Bengal

Dr. Hawz Nwayu

Victoria Global University, UK

Prof. Plewin Amin

Crewe and Alsager College of Higher Education, UK

Dr. (Mrs.) Annifer Zalic

London Guildhall University, London

Dr. (Mrs.) Malin Askiy

Victoria University of Manchester

Dr. ABSALOM

Sixth form College, England

Dr. Nimrod Nivek

London Guildhall University, London

Pressure Drop Measurements in Rectangular Micro-Channel Using Gas Flow

R. Kalaivanan¹ and R. Rathnasamy²

¹ Associate professor, Mechanical Engineering, Annamalai University, India

² Professor, Mechanical Engineering, Annamalai University, India

Abstract: Due to the need for practical cooling technologies which could dissipate high heat fluxes, an experimental study of pressure drop in micro-channel was performed. In this work, laminar flow friction factors were determined using gas (air) as flow medium. Pressure drop vs flow rate data were used to evaluate friction factors in two parallel micro-channels, namely MCP1 and MCP2 (1.0 mm deep x 0.240 mm width and 0.9 mm deep x 0.2 mm width, respectively). Each channel of length 192 mm fabricated on a 304 Stainless Steel substrate by chemical milling Reynolds number was covered between 24 – 5398 for MCP1 and 26 – 6233 for MCP2. Transient pressure drops measured within the channel itself to exclude entrance and exit losses. Friction factor – Reynolds number analyses show that the friction constant are identical as normal channels for gas flow in the laminar region. Transition region lies in $Re > 500$ and transition set off at lower $Re \sim 500$ in comparison to normal channel. The discontinuity in $f - Re$ data identified as transition. Further, it may be possible to identify transition (from laminar region) as the deviation of non-dimensional pressure drop (NDPD) values.

Keywords: Experiments, friction factor, gas, micro-channels, NDPD.

I. Introduction

Rapid advances in microelectronics during the past three decades have brought about a surge of interest in identifying means to reject heat from small surfaces. Heat fluxes in excess of 100 W/cm^2 are often required and in some instance rates that are an order of magnitude higher and desirable. A number of studies have been performed on this topic where the authors concluded that single-phase micro-channel cooling can be used as an effective means of heat rejection [1]. Friction factors measured in laminar and in turbulent flow of gas by [2], using silicon and abraded into glass. It is reported that the friction factors be greater than the classical values, by comparison to the Moody chart for flow in round pipes, not to corresponding charts for rectangular channels. Several investigations ensued which dealt with flow of gases [3]. Fluid flow and heat transfer experiments conducted in rectangular micro-channels by and offered a friction factor correlation ($f=56.9/Re_{Deq}$) close to theoretical value [4]. Based on previous studies and current experimental work was performed to establish transition from $f-Re$ characterisation. In addition, NDPD method is adopted to characterise the same.

II. Fabrication of MCP1 and MCP2

For the present experiments, two test sections were prepared having common features; each channel of length 192 mm were fabricated on a 304 stainless steel substrate of overall dimensions 230 mm x 160 mm x 1.6 mm. This size is chosen

to be comparable to the size of a double Euro PCB so that eventually the results of the study can be applied therein. The two test sections MCP1 and MCP2 were manufactured first by photo chemical etching process as shown in Figure 1. Subsequent to etching of channel the channel header portions were deepened by EDM in order to have negligible pressure loss. MCP1 and MCP2 have 47 and 50 micro-channels of rectangular cross-section 1000 by 240 μm and 900 by 200 μm in width and depth, respectively. Both ends of channels are provided with common header for uniform flow distribution through each channel [5]. These header portions were deepened after etching by EDM in order to have negligible pressure. The channel top covered with another plate of 0.5 mm by vacuum brazing.

III. Experimental set-up and procedure

Figure 2 shows the set-up for the gaseous flow experiments. The experimental set-up for gas flow consists of compressed gas source and a storage reservoir to supply gas to the test section. A pressure regulator is mounted between the gas source and the reservoir to avoid over pressurisation in the circuit under any circumstance. The storage reservoir has capacity of 0.0564 m^3 (56.4 litre). The compressed gas was admitted through the pressure regulator into the storage reservoir to build up a known pressure of about 5 bar. A solenoid valve is connected between the reservoir and test section and it was closed while charging the reservoir. The charged reservoir was allowed to discharge by opening this valve resulting in decrease of the gas pressure in the reservoir. Transient pressure measurements were conducted at a time interval of 5 seconds for both MCP1 and MCP2 at room temperature ($\sim 30^\circ\text{C}$). The reservoir pressure and the pressure drop through the channels between each instant are recorded using differential pressure transducers (make: KELLER Druckmesstechnik; piezoresistive pressure transmitter, PD-23/ 5 bar / 8666.1). The experiment details are given in Table 1.

Table 1. Details of experiments

Fluid	No. of experiments		Re range	
	MCP 1	MCP 2	MCP1	MCP2
Air	16	13	24-5398	26-6233

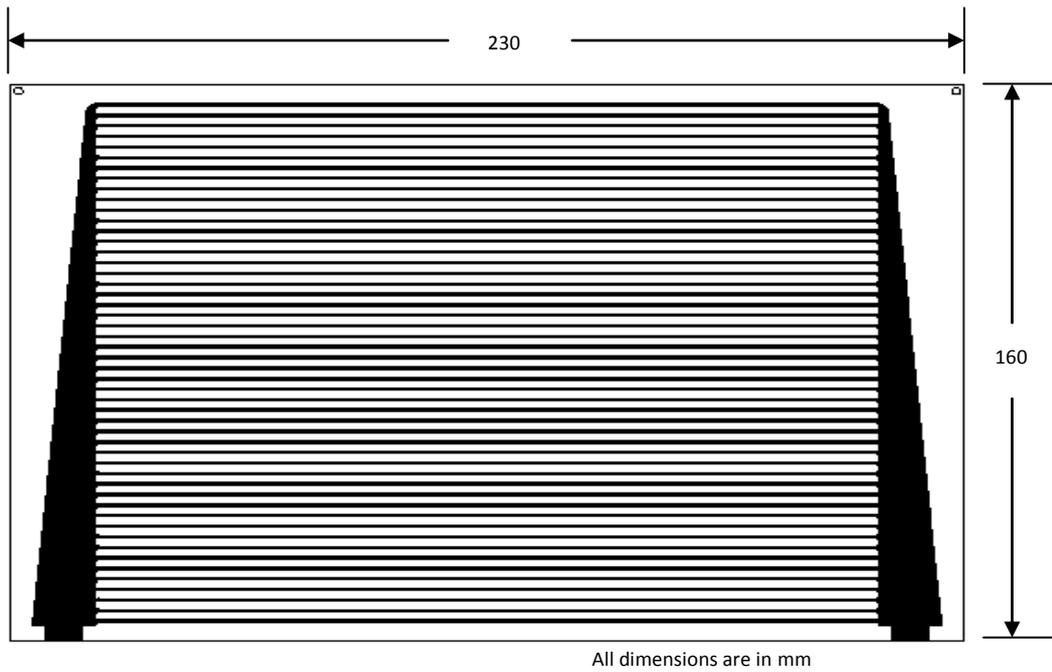


Fig.1 Layout of the parallel type rectangular micro-channel

3.1 Primary data for gas flow

Experiments were conducted for flow of gas as described. The data recorded in the current experimental program are presented graphically. Typical transient pressure curves for gas flow through MCP1 and MCP2 are shown in Figs. 3 and 4 respectively. The transient cylinder pressure and the pressure drop across the test section form the primary data.

The mass flow rate was calculated by numerical differentiation of the reservoir pressure transient. By differentiating the perfect gas relation,

$$pV = m_g RT \quad (1)$$

and the mass flow rate can be calculated as

$$dm_g/d\tau = (dp/d\tau) V/RT \quad (2)$$

The use of perfect gas relation is valid as the compressibility effects under the conditions of experiments are negligible (compressibility factor is almost 1.0).

IV. Gas flow - Data reduction/analysis

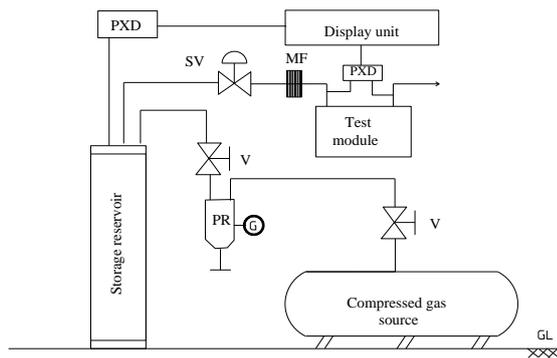


Fig. 2 Gas flow experimental set-up schematic Legend: G- Pressure gauge; MF-Micro filter; PR-Pressure regulator; PXD-Pressure transducer, SV- Solenoid valve V-Valve

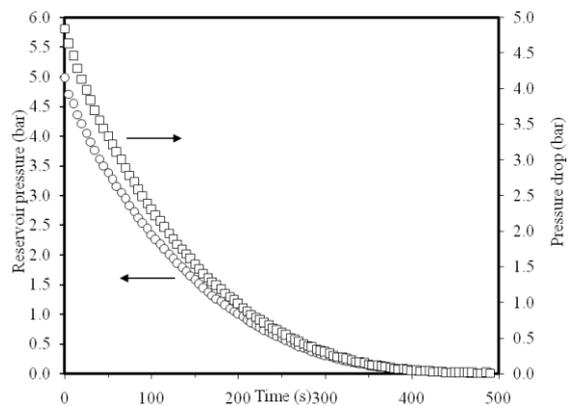


Fig. 3 Transient pressure data for air flow in MCP1

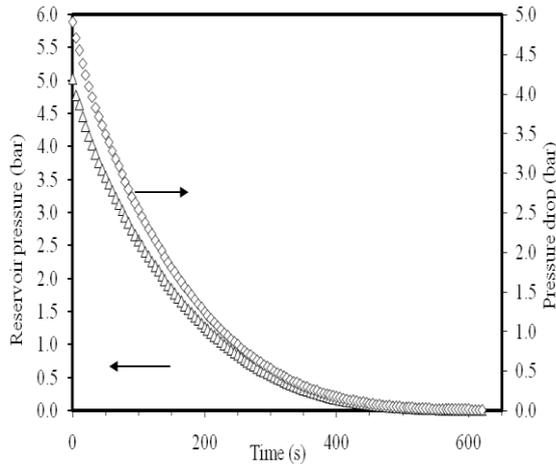


Fig. 4 Transient pressure data for air flow in MCP2

It is assumed that the temperature is constant. The Reynolds number is defined in the conventional way ($\rho v D_{eq} / \mu$) based on cross-sectionally averaged velocity (v) (evaluated using the mass flow rate) and hydraulic equivalent diameter (D_{eq}) which is defined for non-circular duct as follows:

$$D_{eq} = 4WH / (2W + 2H) \quad (3)$$

For the present experiment with gas, Re can be written as follows:

$$Re = 2V(dp/d\tau) / [\mu RT(W + H)] \quad (4)$$

The values of viscosity used [6] is 1.82×10^{-5} Pa-s, for air. The pressure drop across the channel for any time step was taken as the average value at the two instants. The Darcy-Weisbach formula was used for calculating the friction factor as below

$$\Delta p / \rho = f (L / D_{eq}) (v^2 / 2) \quad (5) \text{ where,}$$

Δp is the pressure drop, ρ is the density, f is the friction factor, L / D_{eq} is length to diameter ratio and v is the velocity.

Equation (5) can be rewritten as follows:

$$f = \Delta p \left[\frac{2WH}{W + H} \right]^3 \left(\frac{2\rho}{\mu^2} \right) \frac{Re^{-2}}{L} \quad (6)$$

The density used in the above equation was evaluated at the average pressure between the inlet and outlet of the channel.

V. Gas flow results

The values of friction factor (f) obtained in the present flow experiments are plotted as f vs Re and the line $f = C_{theo} / Re$ have also been drawn in for reference. The C_{theo} is the theoretical value for fully developed laminar flow in a straight channel. It is evident, that there is a deviation between C_{theo} for normal channel assumption and to the present micro-channel. The value of C_{theo} , which depends on the aspect ratio, had been calculated for MCP1 and MCP2 using the data given in F.M.White.

5.1 Friction data

Figures 5 and 6 shows the friction factor vs Reynolds number plots for MCP1 and MCP2 respectively. The laminar flow data were fitted using the following relation,

$$f = C_1 / Re \quad (7)$$

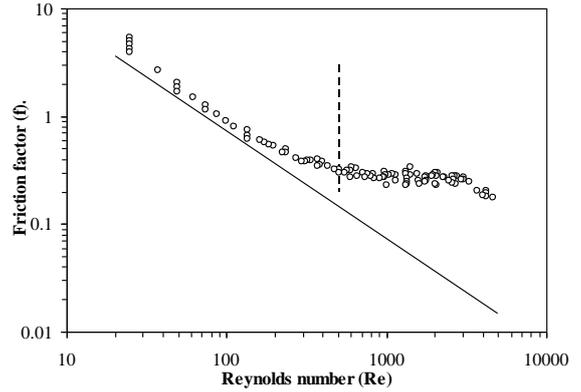


Fig. 5. Plot of f vs Re for flow in MCP1

Legend: O - Experimental, Solid line- Theory, Dashed line- Transition

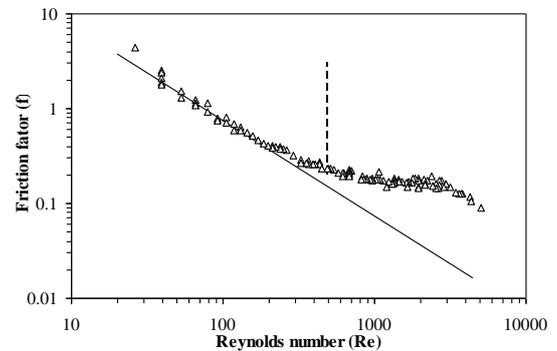


Fig. 6. Plot of f vs Re for flow in MCP2

Legend: Δ - Experimental, Solid line- Theory, Dashed line- Transition

Table 2. Values of C_1 in eq. (7)

Fluid	C_{theo}	Avg e.	mi n.	max .	Std Dev n.	Re_{Tr}^{\dagger}	Remarks on eq. (7)
MCP 1:	73.5						
Air	6	90.38	81.69	98.77	5.12	~500	$Re < 500$
MCP 2:	74.7						
Air	3	91.29	86.47	95.28	2.89	~500	$Re < 500$

[†]Note: Only perceived data are presented.

and the values of C_1 are given in the Table 2. Also Table 2 (Column 2) gives the theoretical values for the two channels. Although, the aspect ratios for MCP1 and MCP2 are of same order, the values of friction constant for MCP1 and MCP2 are more than C_{theo} and are about 22% higher. Comparing the values of MCP1 and MCP2, it is seen that there is a deviation of about 1%.

VI. Non-dimensional pressure drop (NDPD)

An alternative way of presenting friction data is to give a relation between non-dimensional pressure drop ($NDPD=2 \Delta p [(\rho D_{eq}^3)/(\mu^2 L)]$) and Re . Equation (7) re-written becomes

$$NDPD = Re C_1 \quad (8)$$

From the flow measurement data, experimental NDPD was evaluated. Figures 7 and 8 show the plots of NDPD vs Reynolds number for flow of gas in MCP1 and MCP2 respectively. At low Reynolds number values a linear relationship between the NDPD and Reynolds number is evident, wherein experimental NDPD data fall along the solid line (in Fig. 7 for MCP1 and Fig. 8 for MCP2). The solid lines drawn in Figs. 7 and 8 represent the theoretical value of NDPD for laminar flow in straight macro-channels. The dashed lines drawn (in Figs. 7 and 8) represent the experimental value of NDPD based on the average value of C_1 for air (Column 3 in Table 2).

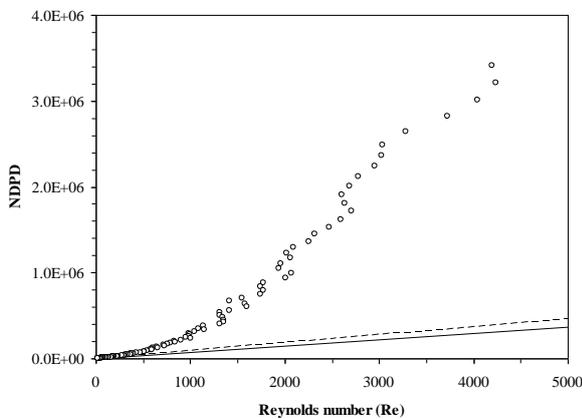


Fig. 7 Plot of NDPD vs Re for flow in MCP1

Legend: O - Experimental, Solid line- C_{theo} , Dashed line- C_1 (exp)

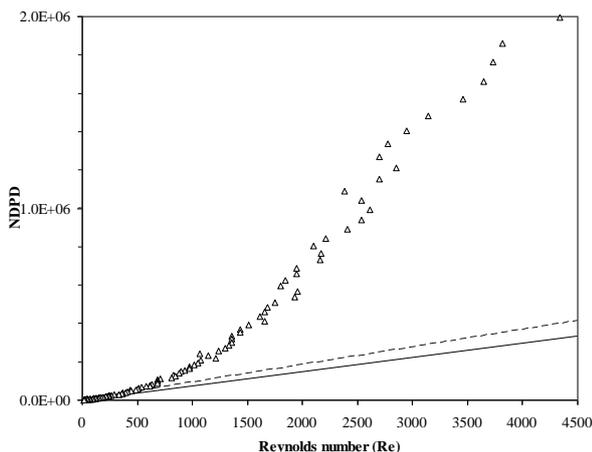


Fig. 8 Plot of NDPD vs Re for flow in MCP2

Legend: Δ - Experimental, Solid line- C_{theo} , Dashed line- C_1 (exp)

Both from the f - Re and NDPD- Re plots, it is clear that in each of the two channels at low Re the pressure drop varies linearly with velocity as would be expected for laminar flow. However, the friction factor is higher than the respective theoretical values for MCP1 and MCP2 (Figs. 5 and 6). It may be mentioned here that other experiments in micro-

channels have shown both higher [2] and lower [8] values of friction factor than those obtained in macro-channels. Further, the Figs. 5 to 8 shown are for one set of experimental data for each fluid and channel.

6.1 Transition Reynolds number (Re_{Tr})

The transition from laminar to turbulent regime is traditionally visualized as a discontinuity in the f vs Re relation. These change(s) of slope as construed to be points of transition as suggested by [7]. Unlike in the conventional flow geometries where one can expect a discontinuity in f vs Re relation, in the case of micro-channels such a sharp discontinuity is absent. Gerlach [7] proposed identification of these mild discontinuities through a plot of normalized pressure drop ($\Delta p D_{eq}^2 \rho / \mu^2$) against Re . This is also analogous to NDPD vs Re (or f vs Re) which has been adopted in the present analysis.

It can be noticed from the f - Re and NDPD- Re plots, for each of the two channels at low Re the pressure drop varies linearly with velocity as would be expected for laminar flow. This feature is the change(s) in slope of NDPD at certain Reynolds number(s) for MCP1 and MCP2 is also evident. These Reynolds numbers in the present analysis are termed as transition Reynolds numbers (Re_{Tr}). For MCP1 and MCP2, large changes in slope occur at $Re \approx 500$. This change in slope appears to be like the characteristic transition to turbulence observed in pipe and channel flows. The derived or identified transition point is marked in Figs. 5 and 6 with vertical dashed line. Further study is required to relate the deviations in the friction curve with different fluids and geometries. However, it is clear that accurate data of friction is required for design of micro-channels.

Three important conclusions arise out of the experiments with gas, namely

- i) The friction factors in MCP1 and MCP2 are the same.
- ii) The friction factors in MCP1 and MCP2 are higher than theory for laminar region.
- iii) The transition was observed in MCP1 and MCP2 at $Re_{Tr} \approx 500$.

VII. Conclusion

Pressure drop, friction factor were measured and calculated for gas flow in rectangular micro-channels. The gas flow experiments have been conducted in MCP1 and MCP2 with Reynolds number range of the order of 24 – 6233, covering laminar and transition flow regimes. The following conclusions emerge from the investigation

1. The transition region in micro-channel flow cannot be described distinctly as in the case of normal channels.
2. Reynolds number alone is inadequate to explain the flow behaviour in micro-channels.
3. To identify the transition, as the deviation of NDPD values or from change in slope.
4. The friction factor values are comparable to conventional tube flow in laminar region.
5. To obtain a generalised correlation for friction factor there is a need for more experiments on various geometries and different gases.

Acknowledgements

The test modules used in this paper are financially supported by Department of Mechanical Engineering, Indian Institute of Science (IISc), Bangalore from Centre for Air Borne Systems project. The authors are grateful to Liquid Propulsion Systems Centre, Indian Space Research Organisation, Bangalore for deploying their expertise in the fabrication of test sections.

References

- [1] M.B. Kleiner, S.A. Kuhn, and K. Habeger, High performance forced air cooling scheme employing micro-channel heat exchangers. *IEEE. Trans. On Components, Packaging and Manufacturing Technology*. Part A 18(4), 1995, 795.
- [2] P. Wu and W.A. Little, Measurement of friction factors for the flow of gases in very fine channels used for Micro-miniature Joule-Thomson Refrigerators. *J. Cryogenics*. 23, 1983, 273.
- [3] G. Hetsroni A. Mosyak, E. Pogrebnyak and L.P. Yarin, Fluid flow in micro-channels. *Int. J. Heat and Mass Transfer*, 52, 2005, 1982-1998.
- [4] J.Y. Jung and H.Y. Kwak, Fluid flow and heat transfer in micro-channels with rectangular cross section. *J. Heat and Mass Transfer*, 44, 2008, 1041-1049.
- [5] R.H. Perry, D. W. Green, and Maloney, *Perry's chemical engineer's hand book*, 6th edition, McGraw-Hill Book Company. 1984.
- [6] C. F. Beaton, and G. F. Hewitt, *Physical property data for the design engineers*. New York Hemisphere publishers, 1989.
- [7] T. Gerlach, Micro-diffusers as dynamic passive valves for micro-pump applications. *Sensors and Actuators*, - A physical 69, 1998, 181-191.
- [8] D. Yu, R. Warrington, R. Barron, and T. Ameen, An experimental and theoretical investigation of fluid flow and heat transfer in micro-tubes. *ASME/ JSME Thermal Engineering Conf. 1*, 1995, 523-530.

A Novel Face Erection and Detection through Fuzzy Grammar

S. Rajaram¹

¹Assistant Professor, Arul College of Technology, Tamilnadu, India

Abstract: This paper presents a structural face construction and detection system. The proposed system consist of the different lightning, rotated facial image, skin color etc. These systems only deal with facial recognition method. The Practical limitations are present in this method. These methods only detect the face when the inputs are full face image. In our method there is no need to show full face of user. We only handle the side view face image. Depending on the face view compare the template face image and fix in which side extract the face up to nose. And construct mirror of the remaining side. Now appropriate face has to be constructed. These formulated face image then compare with original image with local binary pattern.

I. INTRODUCTION

Face recognition has discussed by many researchers. Although human face components are eyes, nose, mouth, skin color and different lighting. Depends on this Components face detection is developed the researches. Among the exacting face detection automation method [1],[2] the full face of input only given to recognition. But the problem is if the image having side view of the face then these systems cannot detect.

The human face having most of meaningful information. Among that the face expression is popular for researchers. There are many method to implement knowledge based system. The location of the face in an image is difficult for face automation system. Feature-based face recognition[3] technique have demonstrated the facial variation. But this process, a large amount of dependency in neighboring pixels. Similarly skin color is not enough to track the face[4].There can be localized. Illumination variations, objects like skin can appear. It can only work well when prior assumptions are satisfied.

In our method we first detect the edges of the face for that image we segment only the face part and remove the hair and ear. Next we check whether input image are left side view or right side view using already stored template. This resulted image primitive extracted up to half model of human face[1] using fuzzy grammar. And mirror of partial result will placed in remaining half appropriate human face. This constructed image can iteration and the resulted image have appropriate face model. This will be recognized with local binary pattern under different lightning condition[2].

II. FACE ERECTION

2.1 EDGE DETECTION

The face construction model block diagram is given below. In this model first we detect edges for facial image. And compare with the template face image

model this will be given a which side view the image are present. To find the face localization [5].In this localization this will measure the various pattern localized. Roberts mask are convolved low pass filter to evaluate horizontal and vertical gradient of the image.

$$I_x = I_{\text{filtered}} \odot [I-1], I_y = I_{\text{filtered}} \odot [I-1] \quad (1)$$

The global threshold is applied for whole input image. So 20% inputs pixels are regarded as edges.

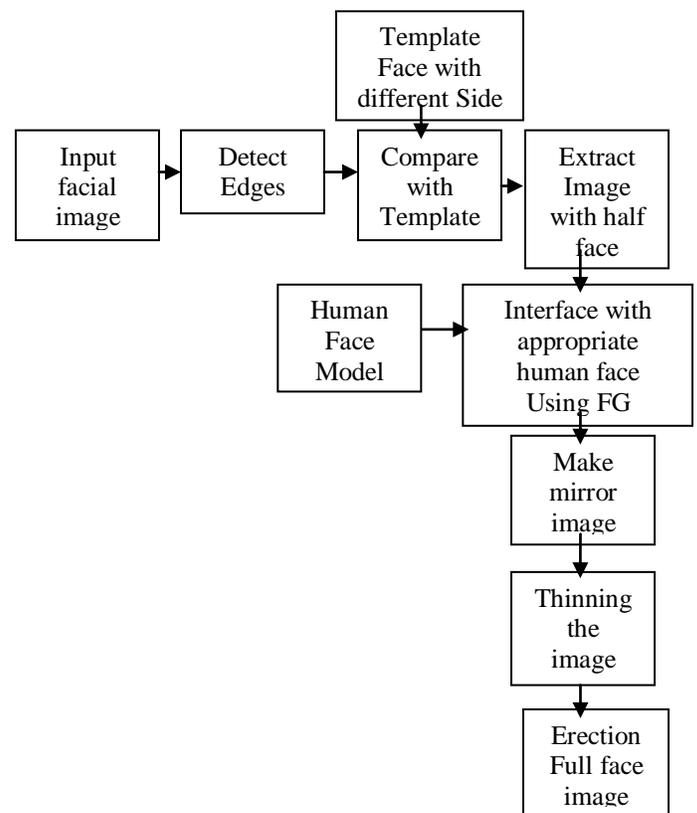


Figure 1. Face erection block diagram.

2.2 EXTRACTION IMAGE USING FUZZY GRAMMAR

Fuzzy Grammar are constructed used in recognition pattern. Fuzzy grammar application are the production rules and the membership values are predefined [1].For our application the human face model has predefined. And just extract the image depends on the face angle variation. A generic algorithm has used for this purpose.

2.3 FUZZY GRAMMARS

A fuzzy grammar (FG) is a 6-tuple (VN, VT, ~S, J, p) where VN is a set of non-terminals, VT is a set of terminals, P is a set of production rules, S is a starting point, J = {pi | i = 1, ..., n, n = cardinality of P} is the set of labels for production rules, and p is a mapping p : J

+ [0,1].

FG generates a fuzzy language (L(FG)) as follows.
 A string $2 E V\$$ is in L(FG) if it is derivable from S, and its grade of membership $p_{\sim}(x) = \max_{l, k} [\min_{i, j} p(r_i)]$ in L(FG) is > 0 , where m is the number of k^{th} derivations that x has in FG; lk is the length of the derivation chain, and r_i is the label of the i^{th} production used in the k^{th} derivation chain, $i = 1, \dots, lk$. If a production $Q \rightarrow \beta$ is visualized as a chain link of strength $p(\sim)$, where T is the label of $Q \rightarrow \beta$, then the strength of a derivation chain is the strength of its weakest link, and therefore $\sim L(FG)(x) = \text{strength of the strongest derivation } S \text{ to } 1$: for all $2 E V$; chain from [1]

In face components extraction there will be used primitive Extraction. Octal chain code use to detect edges of the image. First scan the pixels and find the edges, Octal code is then produced for each edges. Here each pair replaced by digit [1].

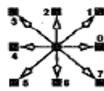


Figure.2. Octal chain Code



Figure.3. Input image and preprocessing result

The extraction primitives for pupil are
 05 7 03 7 0 0 7 0 (0 7) 12 0 6 6 5 4 43 (5 4 4) 3 5 i5 4 4
 5 4 4 5 43 5 410 3 44 (3 4 417 (3 4) 3 (2 4) 3 24 1 1 (0
 113 001 04 io6 1011.[1]

For every face component there was defined fuzzy grammar. Both production and membership values of face components. In face detection, present pattern and noise has to be measured and reject that noise to compare external template..

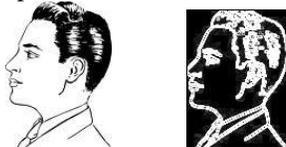


Figure.4. Input side view and edge detected output

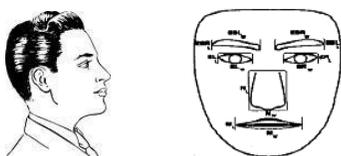


Figure.5. Mirror of input image and human face model.

These input images are used to extract the image with human face model. The pixel are replaced with human model and Extract up to half rang of human face.



Figure.6. Face data and compare with template data



Figure.7. Constructed face

In this stage, the face components are extracted from the constructed face image like eyebrows, eyes, nose, mouth and face edges.

2.4 MERGER

Several overlap occur in face detection. The merger have two step. First position has 3x3 combining filter. The second filter has same work but 20x20 size filter [5].

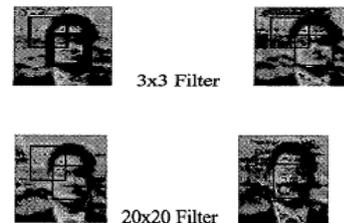


Figure.8 merge the image with 3x3 and 20x20 filter

III. Recognition

We verify the face using facial extraction. First detect the boundary value of the image and extract two region containing eye and eyebrows [3]. The face verification is carry by filtering RFM to extract eye and nose-mouth part. The face locator is first trained up to acceptable error level. And note the important points like eye, nose, mouth. This training method is repeat until the acceptable error level [5].

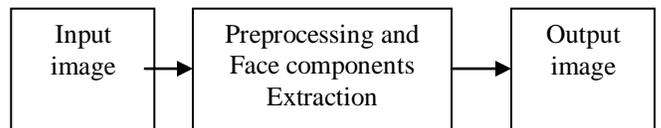


Figure.9. Recognition block diagram.

IV. Experimental Result

In this method we construct a appropriate face model for the given input. This result will testing with windows Microsoft XP on a dual core. The compiler used in IDL 6.3. These will be color and also be grayscale images. In

color image first convert into grayscale and do the face construction. And finally make the recognition. The experimental result show that the implementation of face components extraction stage will consider. in order to overcome consider the light variation and face components variation should be consider.

V. CONCLUSIONS AND FUTURE WORK

A system has been proposed in this paper for facial construction and detection. An input image edges detected first and for that result we fix which side extract the image. The fuzzy Grammar has to use to construct the half face. Next to compare with human face model and construct particular face component. Now the result will appropriate face for particular image. This image then preprocessing and noise are removed then compare with stored image using local binary pattern recognition. The possess about human faces, and makes the final decision. Together with each detected face, a value is produced to denote the degree of membership of the face within the face class.

References

- [1]. A.Z. Kouzani, F. He, and K. Sammut." Constructing a Fuzzy Grammar for Syntactic Face Detection" 0-7803-3280-6/96/%5.00 1996 IEEE
- [2] Satyanadh Gundimada and Vijayan K.Asari." Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces" iee transactions on image processing, vol.18, no.6,june2009

- [3]. Qian CHEN, Haiyuan WU, and Masahiko YACHIDA" Face Detection by Fuzzy Pattern Matching". Proceedings of the Fifth International Conference on Computer Vision (ICCV '95) 0-8186-7042-8/95 \$10.00 © 1995 IEEE
- [4]. Maricor Soriano, Birgitta Martinkauppi, Sami Huovinen and Mika Laaksonen "Skin detection in video under changing illumination conditions" Machine Vision and Media Processing Unit , 0-7695-0750-6/00 \$10.00 0 2000 ieee
- [5]. R.Belaroussi,M.Milgram,L.prevost "Fusion of multiple detectors for face and eye localization" Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis(2005)
- [6]. Stephen Karungaru,Minoru Fukumi and Norio akamatsu "Detection of Human Faces in visual Scenes"

AUTHOR



S. Rajaram received B.E Degree in Electronics and Communication Engineering from Anna University, Chennai India in 2006 He is finish M.Tech degree in Computer and Information Technology from Manonmaniam Sundaranar University in 2011. Currently working as a *Assistant Professor, Arul College of Technology*, His research interests include Signal and Image Processing.

Impact Assessment of Selected Pollution Sources on Groundwater Quality in Wells in Gambari Community, Ogbomosho, Nigeria

Adegbola, Adedayo Ayodele¹, Adewoye, Abosede Olufunmilayo²
Department of Civil Engineering, Ladoko Akintola University of Technology, Ogbomosho, Nigeria¹
Department of Earth Sciences, Ladoko Akintola University of Technology, Ogbomosho, Nigeria²

Abstract : This study focused primarily on the determination of heavy metal properties in water samples from selected hand dug wells in a rapidly urbanized Gambari Community in Ogbomosho, Oyo State, Nigeria. The purpose was to assess the quality of water from these sources. Twenty different water samples were taken and analyzed for physicochemical parameters including pH, electrical conductivity, Cyanide, heavy metals - Cu, Zn, Fe, Pb, CN and Mn. From the results of the analyses, most of the parameters were not found averagely to deviate from the standard or found to be present in high concentrations except for conductivity, pH and Cu. Also, most of the water samples were found to be at least at minimal satisfactory level except for GASP (located within a cassava processing and milling industry) that had ninety five percent (95%) of the parameters tested to be at unsatisfactory level. It is recommended that wells should be sited far away from pollution prone areas. Relevant governmental agencies should, in addition, regularly carry out routine inspection on wells marked for domestic use to safeguard the health of the teeming populace in Gambari Community.

Keywords : Heavy Metals, Physicochemical Parameters, Pollution, Water Samples

I. Introduction

Much of the current concern with regards to environmental protection is focused on water because of its importance in maintaining the human health and health of the ecosystem (Mahananda *et al.*, 2010). Inadequate provision of sufficient volume of drinking water continues to present major problems worldwide to public health (Postel, 1997; Adegbola, 2006). Rapid urbanization of rural areas, industrialization and population growth have been the major causes of stress on the environment leading to human health problems, eutrophication and fish death, coral reef destruction, biodiversity loss, ozone layer depletion and climatic changes (Sadiq, 2002; Bay *et al.*, 2003). Improper disposal of industrial effluents, which is most common in major African urban and rural centres, has led to heavy contamination of the available fresh water resources, reducing the volume of safe water for agricultural, domestic and commercial uses.

Water pollution has been suggested to be the leading worldwide cause of deaths and diseases (Pink, 2006), and that it accounts for the deaths of more than 14,000 people daily (West, 2006). Ground water, according to Ramachandraiah (2004), is the ultimate and most suitable fresh water resource for human consumption in both urban

as well as rural areas. It supports drinking water supply; livestock needs; irrigation; industrial, and many commercial

activities. Groundwater can be contaminated when rainwater carry impurities into streams as runoff or through leaching. Leaching is the downward movement of a dissolved substance through the soil. Also, groundwater is contaminated through the use of fertilizer in farming (Altman and Parizek, 1995). Pesticides and residual fertilizers in the soil, as noted by Adekunle (2009), can leach through the soil along with rainwater. Seepage from effluent bearing water body industrially discharged without proper treatment, also results in the contamination of ground water (Jinwal and Dixit, 2008).

Water of good drinking quality is of basic importance to human physiology and man's continued existence depends very much on its availability (Lamikanra, 1999). The safe portable water is absolutely essential for healthy living. The provision of portable water to the rural and urban population is necessary to prevent health hazards (Nikoladze and Akastal, 1989; Lemo, 2002). Potable water is defined as water that is free from disease producing microorganisms and chemical substances deleterious to health (Ihekoronye and Ngoddy, 1985).

The primary aim of this study is to investigate the impacts of selected pollution sources on the quality of water in hand dug wells in Gambari, Ogbomosho, Nigeria. The study area is underlain by Precambrian rocks of the Nigerian Basement Complex where aquifers are both isolated and compartmentalized. Groundwater localization in this region is controlled by a number of factors which include the parent rock type, the depth, extent and pattern of weathering, thickness of weathered materials, and the degree, frequency and connectivity of fracturing, fissuring and jointing, as well as the type and nature of the fillings in the joint apertures. The study area is characterized by an alternation of two distinct seasons; the rainy season and the dry season. The rainy season lasts for the seven months of April to October with mean annual rainfall of 1000-1400mm, while the dry season begins in November and ends in March.

II. Methodology

A reconnaissance survey was carried out at the study area in order to locate the major sources of pollution. Twenty hand dug wells were randomly selected based on their proximity or closeness to pollution sources. The major pollution sources considered are: cassava mill industry, waste dumpsites and automobile mechanic workshops.

The wells that were sampled were those designated for domestic use and commercial purposes, most especially in preparation of some street vended foods and food canteens. The pH of the samples was measured with a pH

meter (PHEP HANNA 98107) that had been previously calibrated with buffer solution. Electrical conductivity was measured with a conductivity meter (Hitachi 2180) calibrated with potassium chloride.

Heavy metals were determined by digesting a known volume of water sample with HNO₃ (analytical grade). The digested sample was filtered into a 50ml standard flask, made up to the mark with distilled-deionized water. This was stored in a nitric acid prewashed polyethylene bottle in the refrigerator, prior to the instrumental analysis.

The water-extract was analyzed for presence of heavy metals (Pb, Fe, Zn, Cu, CN) by Atomic Absorption Spectrometer. Each sample was analysed in triplicate so as to ascertain the validity of the method, and the average of the results reported. General laboratory quality assurance measures were observed to prevent sample contamination and instrumental errors.

III. Results And Discussion

The pH of a water body is very important in the determination of water quality since it affects other chemical reactions such as solubility and metal toxicity (Fakayode, 2005). The pH of the samples (Figure 1) was between 5.57 (from GA02) and 7.1 (from GASP). Eight (8) of the samples, that is 40%, have pH values below 6.5-8.5 which is the permissible limit of the World Health Organization (WHO), and they are slightly acidic, while the remaining 60% are within the admissible limit, which can be regarded as neutral and unpolluted. The mean value for the potential hydrogen concentration in all the samples analysed is 6.454, which is slightly below the WHO standard.

POTENTIAL HYDROGEN (pH)

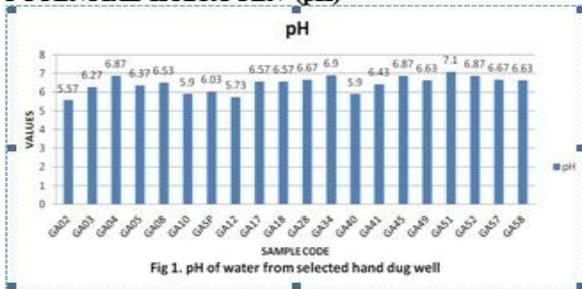


Fig 1. pH of water from selected hand dug well

Electrical conductivity is a measure of water's ability to conduct an electric current and it is related to the amount of dissolved minerals in the water. It does not give an indication of which element is present but higher value of conductivity is a good indicator of the presence of contaminants such as sodium, potassium, chloride or sulphate, and also the amount of total dissolved salts (Orebiyi *et al.*, 2010; Sudhir and Amarjeet, 1999). Conductivity is a good and rapid method to measure the total dissolved ions and is directly related to the total solids in the water sample (Singh *et al.*, 2010). The higher the value of dissolved solids, the greater the amount of ions in water (Bhatt *et al.*, 1999).

Analysis of the results (Figure 2) showed that 50% of the samples have conductivity values above 400 μ S/cm which is the WHO Standard (1995). The range of conductivity of the samples was from 225.67 μ S/cm – 1353

μ S/cm, with the minimum (225.67 μ S/cm) obtained from GA03 and maximum value (1353 μ S/cm) obtained from GASP. High values of conductivity were recorded for GA51, GA12, GA45 and GA17. These indicate a high level of dissolved solids, and subsequently impurities in the water, which can render the water unfit for drinking.

ELECTRICAL CONDUCTIVITY

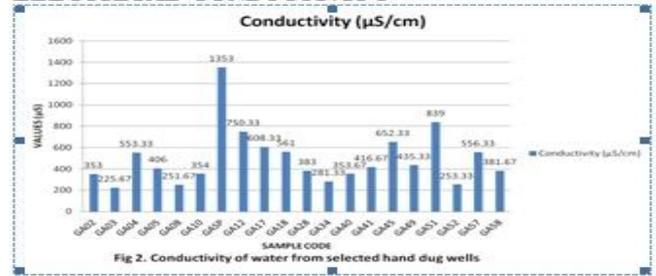


Fig 2. Conductivity of water from selected hand dug wells

Iron is the fourth most abundant element by mass in the earth's crust. In water, it occurs mainly in ferrous or ferric state (Ghulman *et al.*, 2008). It is an essential and non-conservative trace element found in significant concentration in drinking water because of its abundance in the earth's crust. Usually, iron occurring in ground water is in the form of ferric hydroxide, in concentration less than 500 μ g/L (Oyeku and Eludoyin, 2010). Iron is an essential nutrient for erythropoiesis. The shortage of iron causes disease called "anemia" and prolonged consumption of drinking water with high concentration of iron may lead to liver disease called as haemosiderosis (Rajappa *et al.*, 2010; Bhaskar *et al.*, 2010).

In this study, iron content (Figure 3) varied from 0.1mg/l (GA05, GA08, GA17, GA49, GA52) to 0.5mg/l (GASP). Low iron concentration was recorded in 70% of the samples while the remaining 30% recorded high values for concentration of the iron content.

Although, iron is an important dietary requirement in humans needed by hemoglobin and good for several other functions, when high concentrations of iron are absorbed, iron is stored in the pancreas, the liver, the spleen and the heart. This may damage those vital organs. The presence of excess iron in water imparts the taste and it also promotes growth of iron bacteria that hasten rusting process of all the ferrous metals that come in contact with the water (Chukwu *et al.*, 2008).

IRON (Fe²⁺)

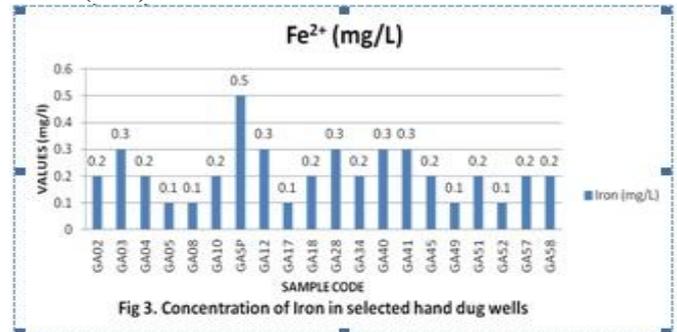


Fig 3. Concentration of Iron in selected hand dug wells

The concentration of manganese (Figure 4) ranged from 0.01mg/l (present in 30% of the samples) to 0.05mg/l (GASP). About 95% of the samples studied have values below the permissible levels for Manganese while the value

recoded for sample collected from GASP was higher compared to the values from others. Excess manganese can interfere with absorption of dietary iron which can result in iron deficiency anemia. It also increase bacterial growth in water. Excessive manganese intake can also cause hypertension in patients older than 40 years.

MANGANESE (Mn²⁺)

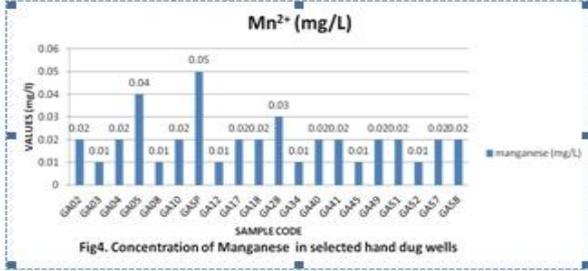


Fig4. Concentration of Manganese in selected hand dug wells

Lead is the most significant of all the heavy metals because it is toxic and harmful even in small amounts (Gregoriadou *et al.*, 2001). Lead enters the human body in many ways. It can be inhaled in dust from lead paints, or waste gases from leaded gasoline. It is found in trace amounts in various foods, notably in fish, which are heavily subjected to industrial pollution. Some old homes may have lead water pipes, which can then contaminate drinking water. Most of the lead taken by humans are removed from the body in urine. As exposure to lead is cumulative over time, there is risk of buildup, particularly in children. Acute effects of lead are inattention, hallucinations and delusions. Poor memory and irritability are symptoms of acute intoxication. Lead absorption in children may affect their development and also results in bone stores of lead (Siddiqui and Sharma, 2009). High concentration of lead in the body can cause death or permanent damage to the central nervous system, the brain, and kidneys (Hanaa *et al.*, 2000). In this study, maximum level of lead concentration (0.03mg/l) was found in water sampled from GASP (Figure 5), and a minimum concentration obtained was below detection level (0mg/l), from water sampled from GA04. The concentration of lead obtained from 95% of the samples are at low level of contamination while the concentration obtained from GASP was found to be slightly higher compared to the rest, indicating the highest level of lead contamination in the sampled wells.

LEAD (Pb²⁺)

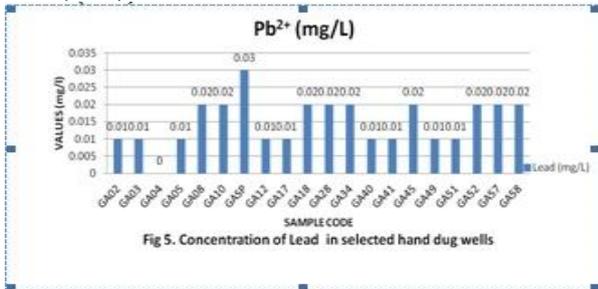


Fig 5. Concentration of Lead in selected hand dug wells

Zinc is one of the important trace elements that play a vital role in the physiological and metabolic process of many organisms. Zinc is a nutritionally essential metal, and its deficiency results in severe health consequences

(Curtis *et al.*, 1996). Nevertheless, higher concentrations of zinc can be toxic to the organism (Rajkovic *et al.*, 2008).

Zinc plays an important role in protein synthesis and is a metal which shows fairly low concentration in surface water due to its restricted mobility from the place of rock weathering or from its natural sources (Rajappa *et al.*, 2010). In this study, a minimum of 0.01mg/l was recorded from five samples: GA08, GA10, GA41, GA49, and GA58 and maximum concentration of 0.05mg/l also recorded from five samples GA04, GASP, GA12, GA45, and GA51. This indicates that the concentration of zinc in the samples (Figure 6) is both at highest and lowest at 23% while the concentration recorded from the remaining 50% ranged from 0.02mg/l to 0.04mg/l.

ZINC (Zn²⁺)

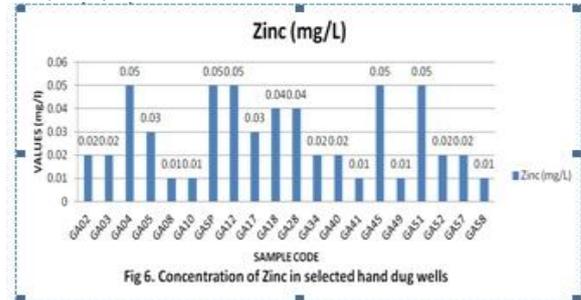


Fig 6. Concentration of Zinc in selected hand dug wells

Copper is an essential component of several enzymes. It is essential for utilization of iron (Curtis *et al.*, 1996). Contamination of drinking water with high level of copper may lead to chronic anemia (Acharya *et al.*, 2008). Copper in excess could impart a bitter taste to water and could promote the corrosion of galvanized iron and steel fittings (Chukwu *et al.*, 2008).

The concentration of copper detected in all the samples (Figure 7) in this study is above the permissible limits of the WHO. The range was from 0.1mg/l (found in 35% of the samples) to 0.3mg/l (found in 15% of the samples). About 50% of the samples have concentration of 0.02mg/l. This indicates that a variation of 0.1mg/l exists within the level of pollution obtained from the samples. GASP and GA28 showed the highest concentration indicating that they have the highest level of pollution from copper.

COPPER (Cu²⁺)

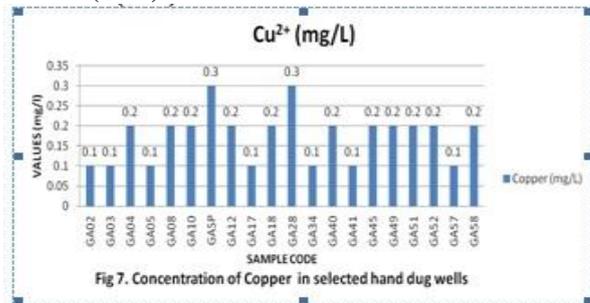


Fig 7. Concentration of Copper in selected hand dug wells

In this study, the concentration of cyanide (Figure 8) obtained ranged from 0.00mg/l (found in 95% of the samples) to 0.7mg/l. Cyanide was only present in water samples collected from GASP, which is located within the cassava processing and milling local industry and is subjected to heavy pollution from cassava waste water

effluent. The concentration of cyanide found is ascertained to be far beyond the required limit.

Therefore, it is expedient that the well in GASP should not be used for both domestic and/or drinking purposes as cyanide is acutely toxic to humans. Relatively low concentrations of cyanide can be toxic to humans as it causes hypoxia, and lactate acidosis which can result in respiratory arrest. Cyanide poisoning also affects organs and systems in the body including the heart.

CYANIDE (CN⁻)

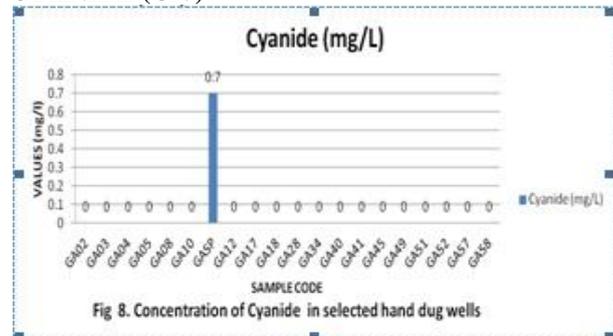


Fig 8. Concentration of Cyanide in selected hand dug wells

IV. Conclusions And Recommendations

The assessment of groundwater quality in the rapidly urbanized Gambari Community, Ogbomoso, Nigeria, was achieved by evaluating the concentrations of certain heavy metals from various locations within the community. From the results of the analyses, most of the parameters were not found averagely to deviate from the standard or found to be present in high concentrations, except for concentration of copper.

Also, most of the water samples were found to be at least at minimal satisfactory level except for GASP (located within a cassava processing and milling industry) that has ninety five percent (95%) of the parameters tested to be at unsatisfactory level.

It is recommended that water from sampled well (GASP) should not be used for domestic purpose in view of high level of contamination. The effects of the presence of heavy metals in the water samples revealed that rapid urbanization coupled with inadequate planning and monitoring are responsible for the pollution of groundwater in Gambari Community.

If the present trend of pollution is not abated, in the near future, consumption of water abstracted from shallow wells will pose a serious health hazard to the generality of the residents in Gambari Community.

In the interim, shallow wells should be sited far away from any visible pollution source. Also, regular inspection and monitoring of existing wells should be prompt, and be on the priority list of the relevant local authority.

References

- [1] Adegbola, A. A. (2006): A Review on Safe Water Provision for Residents of Ibadan City. *Science Focus: International Journal of Biological and Physical Sciences*. 2(1): 139-146.
- [2] Adekunle, A. S. (2009): Effects of Industrial Effluent on Quality of Well Water within Asa Dam Industrial Estate, Ilorin, Nigeria. *Nature and Science*. 7(1): 28-35.
- [3] Altman, S. J. and Parizek, R. R. (1995): Dilution of Non-Point Source Nitrate in Groundwater. *J. Environ. Quality*. 24:707-717.
- [4] Bay, S. M.; Zeng, E. Y.; Lorenson, T. D.; Tran, K. and Alexander, C., (2003). Temporal and Spatial Distributions of Contaminants in Sediments of Santa Monica Bay, California. *Mar. Environ. Res.*, 56 (1-2): 255 – 276.
- [5] Bhaskar, C.V.; Kumar, K. and Nagendrappa, G. (2010): Assessment of Heavy Metals in Water Samples of Certain Locations Situated Around Tumkur, Karnataka, India. Viewed 12 June, 2010, <<http://www.indiaenvironmentportal.org.in/.../Assessment%20of%20heavy%20metals%20in%20water%20samples.pdf>>.
- [6] Bhatt, L. R.; Lacoul, P.; Lekhat, H. D. and Jha P. K. (1999): Physicochemical Characteristics and Phytoplankton of Taudaha Lake, Kathmandu. *Poll. Res.* 18 (14):353-358.
- [7] Chukwu, O.; Mustapha, H. I. and Abdul-Gafar, H. B. (2008): The Effect of Minna Abattoir Waste on Surface Water Quality. I. *Environ. Res. J.*, 2(6):334-338.
- [8] Curtis, D.; Klaassen, C. and Doulls, T. (1996): Toxicology, The Basic Science of Poisons, 5th, International Edition, McGraw Hill, Health Profession Division, USA, 691-721.
- [9] Fakayode, S.O. (2005): Impact Assessment of Industrial Effluent on Water Quality of the Receiving Alaro River in Ibadan, Nigeria. *Ajeam-Ragee* . 10: 1-13.
- [10] Ghulman, B. A.; EL-Bisy, M. S. and Ali, H. (2008): Groundwater Assesment of Makkah Almokarama. *Proceedings of the 12th International Water Technology Conference, Umm Al-Qura University, Makkah*, pp. 1515-1527.
- [11] Gregoriadou, A., Delidou, K.; Dermosonoglou, D.; Tsoumparis, P.; Edipidi, C. and Katsougiannopoulos, B. (2001): Heavy Metals in Drinking Water in Thessaloniki Area, Greece. *Proceedings of the 7th International Conference on Environmental Science and Technology, Aristotle University, Ermoupolis*.
- [12] Hanaa, M.; Eweida, A. and Farag, A. (2000): Heavy Metals in Drinking Water and their Environmental Impact on Human Health. *International Conference on Environmental Hazards Mitigation, Cairo University, Egypt*, pp. 542-556.
- [13] Ihekoronye, A. I.; Ngoddy, P. O. (1985): Integrated Food Sciences and Technology for the Tropics. *Macmillan Press London, Oxford*. pp. 95-195.
- [14] Jinwal, A, and Dixit, S. (2008): Pre and Post Monsoon Variation in Physio-Chemical Characteristic in Groundwater Quality in Bhopal, India. *Asian.j. Exp. Sci.* 22 (3).
- [15] Lamikanra, A. (1999): Essential Microbiology for Students and Practitioner of Pharmacy, Medicine and Microbiology. 2nd ed. Amkra Books Lagos, p. 406.
- [16] Lemo, O. O. (2002): Bacteriology Determination of Water with Long Term Storage. Unpublished B.Sc. Project Report, UNAAB Abeokuta, pp. 40.

- [17] Mahananda,; Mohanty, B. P. and Behera, N. R. (2010): Physico-Chemical Analysis of Surface and Groundwater of Bargarh District, Orissa, India. *IJRRAS* 2 (3).
- [18] Nikoladze, G. D. M. and Akastal, S. (1989): Water Treatment for Public and Industrial Supply. *MIR Publisher Moscou*, p. 163.
- [19] Orebiyi, E. O.; Awomeso, J. A.; Idowu, O. A.; Martins, O.; Oguntoke, A. and Taiwo, A. M. (2010): Assessment of Pollution Hazards of Shallow Well Water in Abeokuta and Environs, Southwest, Nigeria. *American Journal of Environmental Science*, 6(1):50-56.
- [20] Oyeku, O. T. and Eludoyin, A. O. (2010): Heavy Metal Contamination of Groundwater Resources in a Nigerian Urban Settlement. *African Journal of Environmental Science and Technology*, 4(4):201-214.
- [21] Pink, D. H. (2006): Investing in Tomorrow's Liquid Gold. *Yahoo*. <http://finance.yahoo.com/columnist/article/trenddesk/3748>.
- [22] Postel, S. (1997): Last Oasis: Facing Water Security. New York, Norton, p 17-191.
- [23] Rajappa,; B.; Manjappa, S. and Puttaiah, E.T. (2010): Monitoring of Heavy Metal Concentration in Groundwater of Hakinaka Taluk, India. *Contemporary Engineering Sciences*, 3(4):183-190.
- [24] Rajkovic, M. B.; Lacnjevac, C. M.; Ralevic, N. R.; Stojanovic, M. D.; Toskovic, D. V.; Pantelic, G. K.; Ristic, N. M. and Jovanic, S. (2008): Identification of Metals (Heavy and Radioactive) in drinking Water by Indirect Analysis Method Based on Scale Tests. *Sensors*, 8:2188-2207.
- [25] Ramachandraiah, C. (2004): Right to Drinking Water in India., *Centre for Economic and Social Studies*. 56.
- [26] Sadiq, M., (2002): Metal Contamination in Sediments from a Desalination Plant Effluent Outfall Area. *Sci. Total Environ.*, 287 (1), 37-44.
- [27] Singh M. R.; Gupta, A.; Beeteswari, K. H. (2010): Physico-Chemical Properties of Water Samples from Manipur River System, India. *J. Appl. Sci. Environ. Vol. 14 (4) 85 – 89*.
- [28] Sudhir, D. and Amarjeet, K. (1999): Physicochemical Characteristics of Underground Water in Rural Areas of Tosham Subdivisions, Bhiwani District, Haryana, *J. Environ Poll.*, 6 (4), 281.
- [29] West, L. (2006): World Water Day: A Billion People Worldwide Lack Safe Drinking Water. In: <http://environment.about.com/od/environmentalevents/a/waterdayqa.htm>
- [30] World Health Organisation (1995) Guidelines for Drinking Water Quality, W.H.O. Geneva.

Unsteady MHD free convective flow past a semi-infinite vertical porous plate

K. Rajasekhar¹, G. V. Ramana Reddy², B. D. C. N. Prasad³

1. Dept. of Mathematics, RVR & JC College of Engineering, Guntur, (India) -522 019

2. Dept. of Mathematics, Usha Rama College of Engineering and Technology, Telaprolu, Krishna (India)-521109.

3. Dept. of Mathematics, Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada, (India)-520007.

Abstract : In this article, we studied the effects of variable viscosity and thermal conductivity on an unsteady two-dimensional laminar flow of a viscous incompressible electrically conducting fluid past a semi-infinite vertical plate taking into account the mass transfer. The fluid viscosity is assumed to vary as a linear function of temperature. The governing equations for the flow are transformed into a system of non-linear ordinary differential equations are solved by a perturbation technique. The effects of the various parameters on the velocity, temperature, concentration and skin-friction profiles are presented graphically and discussed qualitatively.

Keywords: Radiation parameter, Thermal conductivity, MHD, porous medium and viscosity.

I. Introduction

Natural convection flow over vertical surfaces immersed in porous media has paramount importance because of its potential applications in soil physics, geohydrology, and filtration of solids from liquids, chemical engineering and biological systems. Study of fluid flow in porous medium is based upon the empirically determined Darcy's law. Such flows are considered to be useful in diminishing the free convection, which would otherwise occur intensely on a vertical heated surface. In addition, recent developments in modern technology have intensified more interest of many researchers in studies of heat and mass transfer in fluids due to its wide applications in geothermal and oil reservoir engineering as well as other geo-physical and astrophysical studies.

Cramer, K. R. and Pai, S. I. [1] taken transverse applied magnetic field and magnetic Reynolds number are assumed to be very small, so that the induced magnetic field is negligible Muthucumaraswamy et al. [2] have studied the effect of homogenous chemical reaction of first order and free convection on the oscillating infinite vertical plate with variable temperature and mass diffusion. Sharma [3] investigate the effect of periodic heat and mass transfer on the unsteady free convection flow past a vertical flat plate in slipflow regime when suction velocity oscillates in time. Chaudhary and Jha [4] studied the effects of chemical reactions on MHD micropolar fluid flow past a vertical plate in slip-flow regime. Anjalidevi et al. [5] have examined the effect of chemical reaction on the flow in the presence of heat transfer and magnetic field. Muthucumaraswamy et al. [6] have investigated the effect of thermal radiation effects on flow past an impulsively started infinite isothermal vertical plate in the presence of first order chemical reaction. Moreover, Al-Odat and Al-

Azab [7] studied the influence of magnetic field on unsteady free convective heat and mass transfer flow along an impulsively started semi-infinite vertical plate taking into account a homogeneous chemical reaction of first order.. The effect of radiation on the heat and fluid flow over an unsteady stretching surface has been analyzed by El-Aziz [8]. Singh et. al. [9] studied the heat transfer over stretching surface in porous media with transverse magnetic field. Singh et. al. [10] and [11] also investigated MHD oblique stagnation-point flow towards a stretching sheet with heat transfer for steady and unsteady cases. Elbashbeshy et. al. [12] investigated the effects of thermal radiation and magnetic field on unsteady boundary layer mixed convection flow and heat transfer problem from a vertical porous stretching surface. Ahmed Sahin studied influence of chemical reaction on transient MHD free Convective flow over a vertical plate. Recently, the chemical reaction, heat and mass transfer on MHD flow over a vertical stretching surface with heat source and thermal stratification have been presented by Kandasamy et al.[13]. The opposing buoyancy effects on simultaneous heat and mass transfer by natural convection in a fluid saturated porous medium investigated by Angirasa et al.[14]. Ahmed [15] investigates the effects of unsteady free convective MHD flow through a porous medium bounded by an infinite vertical porous plate. Ahmed Sahin [16] studied the Magneto hydrodynamic and chemical reaction effects on unsteady flow, heat and mass transfer characteristics in a viscous, incompressible and electrically conducting fluid over a semi-infinite vertical porous plate in a slip-flow regime.

The objective of the present study is to investigate the effect of various parameters like chemical Reaction parameter, thermal Grashof number, mass Grashof number, magnetic field parameter, radiation parameter on convective heat transfer along an semi-infinite vertical plate in porous medium. The governing non-linear partial differential equations are first transformed into a dimensionless form and thus resulting non-similar set of equations has been solved using the perturbation technique. Results are presented graphically and discussed quantitatively for parameter values of practical interest from physical point of view.

II. Mathematical Formulation

In a situation of two dimensional unsteady laminar natural convection flows of a viscous, incompressible, electrically conducting, radiating fluid past an impulsively started semi-infinite vertical plate in the presence of transverse magnetic field with viscous dissipation is considered. The fluid is assumed to be gray, absorbing-

emitting but non-scattering. The x -axis is taken along the plate in the upward direction and the y -axis is taken normal to it. The fluid is assumed to be slightly conducting and hence the magnetic Reynolds number is much less than unity and the induced magnetic field is negligible in comparison with the transverse applied magnetic field. Initially, it is assumed that the plate and the fluid are at the same temperature T'_∞ and concentration level C'_∞ everywhere in the fluid. At time $t' > 0$, the plate starts moving impulsively in the vertical direction with constant velocity u_0 against the gravitational field. Also, the temperature of the plate and the concentration level near the plate are raised to T'_w and C'_w , respectively and are maintained constantly thereafter. It is assumed that the concentration C' of the diffusing species in the binary mixture is very less in the comparison to the other chemical species, which are present and hence the Soret and Dufour effects are negligible. It is also assumed that there is no chemical reaction between the diffusing species and the fluid. Then, under the above assumptions, in the absence of an input electric field, the governing boundary layer equations with Boussinesq's approximation are

Continuous equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0 \quad (1)$$

Momentum conservation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = g\beta(T - T_\infty) + g\beta^*(C - C_\infty) + v \frac{\partial^2 u}{\partial y^2} - \frac{\sigma B_0^2}{\rho} u - \frac{v}{K'} u \quad (2)$$

Energy conservation

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \alpha \frac{\partial^2 T}{\partial y^2} - \frac{1}{\rho c_p} \frac{\partial q_r}{\partial y} + \frac{v}{c_p} \left(\frac{\partial u}{\partial y} \right)^2 \quad (3)$$

Species conservation

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} = D \frac{\partial^2 C}{\partial y^2} - K'_r (C - C_\infty) \quad (4)$$

The initial and boundary conditions are as follows:

$$\left. \begin{aligned} t \leq 0, u = 0, v = 0, T = T_\infty, C = C_\infty \quad \forall y \\ t > 0, u = u_0, v = 0, T = T_w, C = C_w \quad \text{at } y = 0 \\ u \rightarrow 0, T \rightarrow T_\infty, C' \rightarrow C_\infty \quad \text{as } y \rightarrow \infty \end{aligned} \right\} \quad (5)$$

Thermal radiation is assumed to be present in the form of a unidirectional flux in the y -direction i.e., q_r (transverse to the vertical surface). By using the Rosseland approximation, the radiative heat flux q_r is given by

$$q_r = -\frac{4\sigma_s}{3k_c} \frac{\partial T'^4}{\partial y} \quad (6)$$

where σ_s is the Stefan-Boltzmann constant and K_c - the mean absorption coefficient. In the Rosseland approximation, the present analysis is limited to optically thick fluids. If temperature differences within the flow are sufficiently small, then equation (6) can be linearized by expanding T'^4 into the Taylor series about T'_∞ , which after neglecting higher order terms takes the form:

$$T'^4 \cong 4T'^3_\infty T' - 3T'^4_\infty \quad (7)$$

In view of equations (6) and (7), equation (3) reduces to

$$\frac{\partial T'}{\partial t} + u \frac{\partial T'}{\partial x} + v \frac{\partial T'}{\partial y} = \alpha \frac{\partial^2 T'}{\partial y^2} + \frac{16\sigma_s T'^3_\infty}{3k_c \sigma c_p} \frac{\partial^2 T'}{\partial y^2} + \frac{v}{c_p} \left(\frac{\partial u}{\partial y} \right)^2 \quad (8)$$

The non-dimensional quantities introduced in these equations are defined as

$$\left. \begin{aligned} X = \frac{xu_0}{v}, Y = \frac{yu_0}{v}, t = \frac{t'u_0^2}{v}, K_r = \frac{K'_r v}{u_0^2}, U = \frac{u}{u_0}, \\ V = \frac{v}{u_0}, r = \frac{v g \beta (T'_w - T'_\infty)}{u_0^3}, Gm = \frac{v g \beta^* (C'_w - C'_\infty)}{u_0^3}, \\ N = \frac{k_c k}{4\sigma_s T'^3_\infty}, M = \frac{\sigma \beta_0^2 v}{u_0^2}, T = \frac{T' - T'_\infty}{T'_w - T'_\infty}, C = \frac{C' - C'_\infty}{C'_w - C'_\infty}, \\ K = \frac{K' u_0^2}{v^2}, Pr = \frac{v}{\alpha}, Sc = \frac{v}{D} \end{aligned} \right\} \quad (9)$$

In a situation where only one dimensional flow is considered, the above set of equations (1), (2), (8) and (4) are reduced to the following non-dimensional form:

$$\frac{\partial V}{\partial Y} = 0 \Rightarrow V = -V_0 \quad (\text{where } V_0 = 1) \quad (10)$$

$$\frac{\partial U}{\partial t} - \frac{\partial U}{\partial Y} = GrT + GmC + \frac{\partial^2 U}{\partial Y^2} - \left(M + \frac{1}{K} \right) U \quad (11)$$

$$\frac{\partial T}{\partial t} - \frac{\partial T}{\partial Y} = \frac{1}{Pr} \left(1 + \frac{4}{3N} \right) \frac{\partial^2 T}{\partial Y^2} \quad (12)$$

$$\frac{\partial C}{\partial t} - \frac{\partial C}{\partial Y} = \frac{1}{Sc} \frac{\partial^2 C}{\partial Y^2} - K_r C \quad (13)$$

The corresponding initial and boundary conditions are as follows

$$\left. \begin{aligned} t \leq 0: U = 0, T = 0, C = 0 \quad \forall y \\ t > 0: U = 1, T = 1, C = 1 \quad \text{at } Y = 0 \\ U \rightarrow 0, T \rightarrow 0, C \rightarrow 0 \quad \text{as } Y \rightarrow \infty \end{aligned} \right\} \quad (14)$$

III. Solution Of The Problem

In order to solve equations (11) - (13) with respect to the boundary conditions (14) for the flow, let us take

$$U(y, t) = U_0(y) + U_1(y) e^{\omega t} \quad (15)$$

$$T(y, t) = T_0(y) + T_1(y) e^{\omega t} \quad (16)$$

$$C(y, t) = C_0(y) + C_1(y) e^{\omega t} \quad (17)$$

Substituting the Equations (15) - (17) in Equations (11) - (13), we obtain:

$$u''_0 + u'_0 - (M + 1/K) u_0 = -[GrT_0 + GmC_0] \quad (18)$$

$$u''_1 + u'_1 - (M + 1/K + \omega) u_1 = -[GrT_1 + GmC_1] \quad (19)$$

$$T''_0 + K_1 T'_0 = 0 \quad (20)$$

$$T''_1 + K_1 T'_1 - \omega K_1 T_1 = 0 \quad (21)$$

$$C''_0 + ScC'_0 - ScK_r C_0 = 0 \quad (22)$$

$$C''_1 + ScC'_1 - Sc(K_r + \omega) C_1 = 0 \quad (23)$$

where prime denotes ordinary differentiation with respect to y .

The corresponding boundary conditions can be written as

$$u_0 = 0, T_0 = 1, T_1 = 0, C_0 = 1, C_1 = 0 \quad \text{at } y = 0 \quad (24)$$

$$u_0 = 0, T_0 = T_1 = 0, C_0 = C_1 = 0 \quad \text{as } y \rightarrow \infty$$

Solving equations (19) – (23) under the boundary conditions (24), we obtain the velocity, temperature and concentration distribution in the boundary layer as:

$$U(y, t) = (1 + A_3)e^{-m_3 y} - A_1 e^{-m_2 y} - A_2 e^{-m_1 y}$$

$$T(y, t) = e^{-m_2 y}$$

$$C(y, t) = e^{-m_1 y}$$

where

$$m_1 = 0.5 \left[Sc + \sqrt{Sc^2 + 4K_r Sc} \right],$$

$$n_1 = 1 + 4/3N, m_2 = \sqrt{Pr/n_1};$$

$$A_1 = Gr / [m_2^2 - m_2 - (M + 1/K)],$$

$$A_2 = Gm / [m_1^2 - m_1 - (M + 1/K)]$$

$$A_3 = A_1 + A_2; K_1 = \frac{Pr}{1 + 4/3N}$$

Skin-friction: The dimensionless shearing stress on the surface of a body, due to the fluid motion, is known as skin-friction and is defined by the Newton's law of viscosity.

The skin-friction is

$$\tau = \left(\frac{\partial U}{\partial y} \right)_{y=0} = -m_3(1 + A_3) + A_1 m_2 + A_2 m_1$$

IV. Results And Discussion

In order to get a physical insight in to the problem the effects of various governing parameters on the physical quantities are computed and represented in Figures 1-13 and discussed in detail. The formulation of the effects of chemical reaction and heat absorption on MHD convective flow and mass transfer of an incompressible, viscous fluid along a semi infinite vertical porous moving plate in a porous medium has been performed in the preceding sections. This enables us to carry out the numerical calculations for the distribution of the velocity, temperature and concentration across the boundary layer for various values of the parameters.

The influence of Magnetic field on the velocity profiles has been studied in Fig .1. It is seen that the increase in the applied magnetic intensity contributes to the decrease in the velocity. Further, it is seen that the magnetic influence does not contribute significantly as we move away from the bounding surface. The influence of the porosity of the boundary on the velocity of the fluid medium has been shown in Fig 2. It is seen that as the porosity of the fluid bed increases, the velocity also increases which is in tune with the realistic situation. Further, the porosity of the boundary does not influence of the fluid motion as we move far away from the bounding surface. The contribution of radiation parameter on the velocity profiles is noticed in Fig.3. It is observed that the radiation parameter increases, the velocity field is an increasing. Further, it is noticed that the velocity decreases as we move away from the plate which is found to be independent of radiation parameter. The effect of Prandtl number on the velocity profiles has been illustrated in Fig.4. It is observed that as the Prandtl number increases, the velocity decreases in general. The dispersion in the velocity profiles is found to be more significant for smaller

values of Pr and not that significant at higher values of Prandtl number.

The influence of Schmidt number Sc on velocity profiles has been illustrated in Fig.5. It is observed that, while all other participating parameters are held constant and Sc is increased, it is seen that the velocity decreases in general. Further, it is noticed that as we move far away from the plate, the fluid velocity goes down. The effect of Grashof number on the velocity profiles as shown in Fig. 6. Increase in Gr contributes to an increase in velocity when all other parameters that appear in the velocity field are held constant. Also it is noticed that as we move away from the plate the influence of Gr is not that significant. The effect of modified Grashof number Gc on the velocity profiles is observed in Fig.7. Increase in Gc is found to influence the velocity to increase. Also, it is seen that as we move far away from the plate it is seen that the effect of Gc is found to be not that significant. Fig.8 shows that the effect of increasing the chemical reaction parameter on velocity profiles. It is noticed that velocity of flow field are decreasing, as the values of chemical reaction are increasing. The Effect of Prandtl number on the temperature field has been illustrated in Fig.9. It is observed that as the Prandtl number increases, the temperature in the fluid medium decreases. Also, as we move away from the boundary, the Prandtl number has not much of significant influence on the temperature. The dispersion is not found to be significant. Fig.10 illustrates the influence of the radiation parameter on the temperature profiles in the boundary layer. As radiation parameter increases, temperature distributions increase when the other physical parameters are fixed.

The influence of Schmidt number on the concentration is illustrated in Fig. 11. It is observed that increase in Sc contributes to decrease of concentration of the fluid medium. Further, it is seen that Sc does not contributes much to the concentration field as we move far away from the bounding surface. Fig.12 shows that the effect of increasing the chemical reaction parameter on concentration profiles. It is noticed that species concentration are decreasing, as the values of chemical reaction are increasing. The effect of chemical reaction on velocity and temperature is less dominant in comparison to concentration. Skin friction for various values of magnetic field strength is portrayed through Fig 13. It is seen that skin friction decreases, as magnetic parameter increases, whereas it is increasing when radiation parameter is increasing.

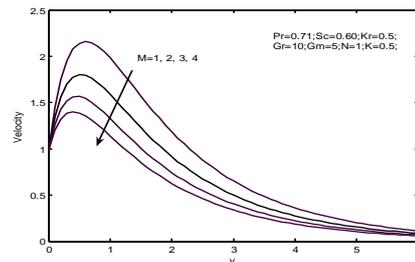


Fig. 1. Effects of magnetic parameter on velocity profiles.

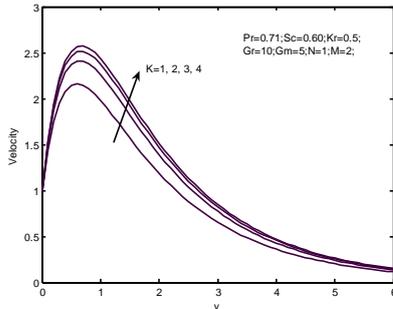


Fig.2. Effects of permeability parameter on velocity profiles.

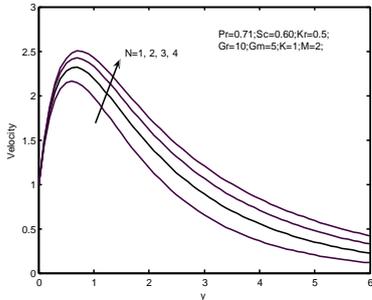


Fig.3. Effects of radiation parameter on velocity profiles.

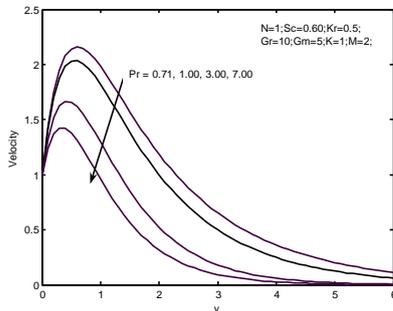


Fig.4. Effects of Prandtl number on velocity profiles.

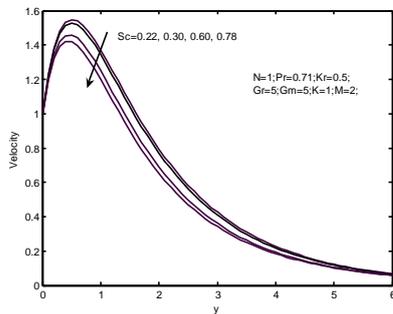


Fig.5. Effects of Schmidt number on velocity profiles.

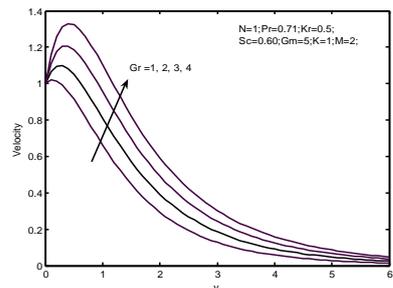


Fig.6. Effects of Grashof number on velocity profiles.

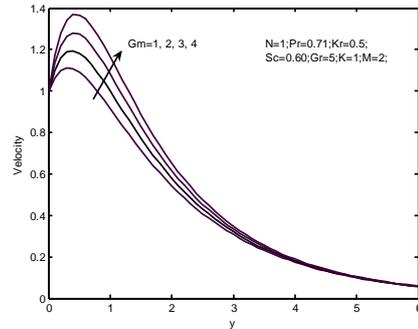


Fig.7. Effects of modified Grashof number on velocity profiles.

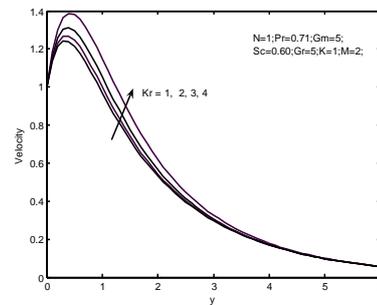


Fig.8. Effects of magnetic parameter on velocity profiles.

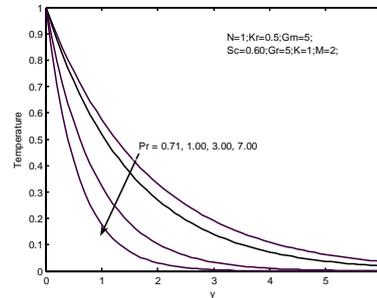


Fig.9. Effects of prandtl number on temperature profiles.

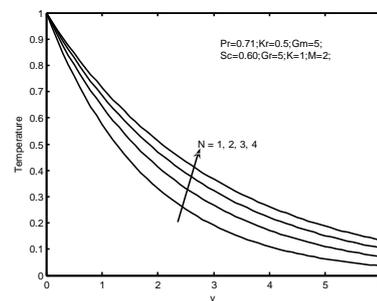


Fig.10. Effects of radiation parameter on temperature profiles.

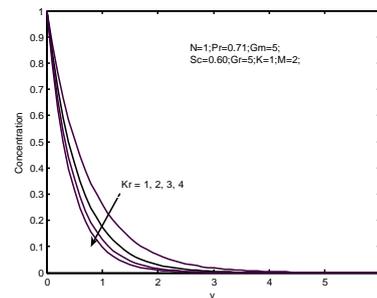


Fig.11. Effects of chemical reaction parameter on concentration profiles

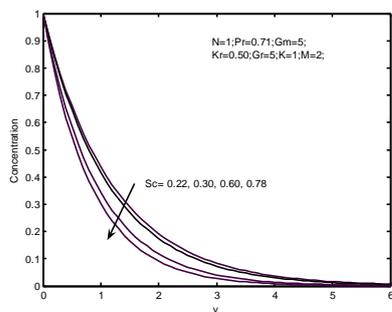


Fig.12. Effects of Schmidt number on concentration profiles.

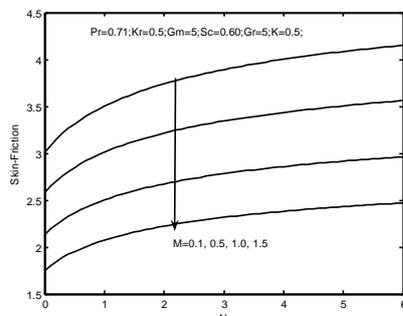


Fig.13. Effects of magnetic parameter on skin-friction.

References

- [1] Cramer, K. R. and Pai, S. I. Magneto fluid Dynamics for Engineering and Applied Physicists (Mc Graw Hill, New York), (1973).
- [2] Muthucumaraswamy.R and Meenakshisundaram.S. Theoretical study of chemical reaction effects on vertical oscillating plate with variable temperature. Theoret. Appl. Mech., vol.33, No.3 (2006), pp.245-257.
- [3] Sharma, P. K. Fluctuating thermal and mass diffusion on unsteady free convective flow past a vertical plate in slip-flow regime, Latin American Applied Research, 35, 313-319,(2005).
- [4] Anjalidevi, S.P. and R. Kandasamy. Effects of a chemical reaction heat and mass transfer on MHD flow past a semi infinite plate. Z.Angew. Math. Mech.,(2000) 80:697-701.
- [5] Chaudhary, R. C. and Jha, A. K.,Effects of chemical reactions on MHD micropolar fluid flow past a vertical plate in slip-flow regime, Applied Mathematics and Mechanics, 29 (9), 1179-1194, (2008).
- [6] Muthucumaraswamy, R. and Chandrakala, P. Radiation heat and mass transfer effects on moving isothermal vertical plate in the presence of chemical reaction. J. of Applied Mechanics and Engineering, vol. 11 NO.3, (2006), pp 639-646.
- [7] Al-Odat, M. Q. and Al-Azab. Influence of chemical reaction on transient MHD free convection over a moving vertical plate, Emirates J. Engg. Res., 12 (3), 15-21. (2007).
- [8] M.A. El-Aziz, "Radiation effect on the flow and heat transfer over an unsteady stretching sheet", International Communications in Heat and Mass Transfer, Vol. 36, pp 521-524, (2009).
- [9] P. Singh, N.S. Tomer and D. Sinha, "Numerical study of heat transfer over stretching surface in porous media with transverse magnetic field", Proceeding of International Conference on Challenges and application of Mathematics in Sciences and Technology (2010), ISBN 023- 032-875-X, pp 422-430.
- [10] P. Singh, N.S. Tomer, S. Kumar and D. Sinha, "MHD oblique stagnation-point flow towards a stretching sheet with heat transfer", International Journal of Applied Mathematics and Mechanics, Vol. 6, no.13, pp 94-111, (2010).
- [11] P. Singh, A. Jangid, N.S. Tomer, S. Kumar and D. Sinha, "Effects of Thermal Radiation and Magnetic Field on Unsteady Stretching Permeable Sheet in Presence of Free Stream Velocity", International Journal of Information and Mathematical Sciences, Vol.6, no.3, pp-63-69, (2010).
- [12] E. M. A. Elbashaeshy, D. M. Yassmin and A. A. Dalia, "Heat Transfer Over an Unsteady Porousv Stretching Surface Embedded in a Porous Medium with Variable Heat Flux in the Presence of Heat Source or Sink", African Journal of Mathematics and Computer Science Research Vol. 3, no.5, pp 68-73, (2010).
- [13] Kandasamy, R., Periasamy, K. and Sivagnana Prabhu, K.K. Chemical reaction, heat and mass transfer on MHD flow over a vertical stretching surface with heat source and thermal stratification effects, Int. J. Heat and Mass Transfer, 48 (21-22), 4557-4561,(2005).
- [14] Angirasa, D., Peterson, G. P. and Pop, I. Combined heat and mass transfer by natural convection with opposing buoyancy effects in a fluid saturated porous medium, Int. J. Heat Mass Trans., 40(12), 2755-2773,(1997).
- [15] Ahmed, S. Effects of unsteady free convective MHD flow through a porous medium bounded by an infinite vertical porous plate, Bull. Cal. Math. Soc., 99 (5), 511-522,(2007).
- [16] Ahmed Sahin. Influence of chemical reaction on transient mhd free Convective flow over a vertical plate in slip-flow Regime. Emirates Journal for Engineering Research, 15 (1), 25-34 (2010).

Capturing Web Log and Performing Preprocessing of the User's Accessing Distance Education System

Mr. Shivkumar Khosla¹, Mrs. Varunakshi Bhojane²
^{1,2} (Department of Computer Engineering, Mumbai University, India)

ABSTRACT: In this paper, we have introduced a concept of capturing different web log file, while the user is accessing the Distance Education System website. Web log file can be further used in pattern discovery and pattern analysis process. Web log file is saved in text (.txt) format with "comma" separated attributes. Log files can't be directly used for pattern discovery process because it consists of irrelevant and inconsistent access information. Therefore there is need of Web log preprocessing which includes different techniques such as field extraction, data cleaning, data filtering, and data summarization. We have discussed different types of web log files and preprocessing techniques.

Keywords: Data Preprocessing, Web Log Mining, Web Log File, Web Personalization, and Web Usage Mining.

I. Introduction

With the rapid development of World Wide Web (WWW), Web application is increasing at enormous speed and its users are increasing at exponential speed. Distance Education System has been widely used in the field of education. Students try to learn from web so, there is a need to provide students with the personalize teaching in Distance Education System which help them to understand the concept in much more better way than the traditional distance education system.

Web Log mining provides a log or access information of the users accessing the Distance Education System. Web Log Mining is an effective and mostly used technique of Web Usage Mining. Most Application performs Web Log Mining to improve web personalization, future prediction and also increasing performance in terms of response from the Web server.

1. Web Log Content

There are three main sources to get the web log file [Fig 1.] such as [2]

1. Client Log File
2. Proxy Log File
3. Server Log File

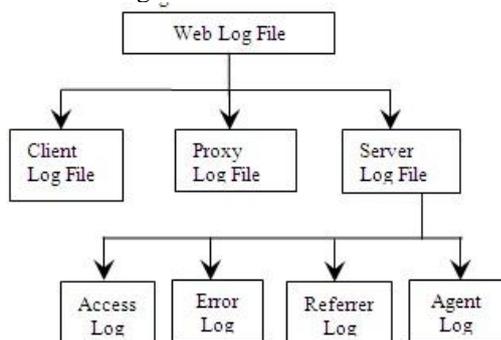


Fig.1 Types of Web log File

Client Log Files mostly consist of authentic i.e. username and password but it does not consist of any of the browser information.

Proxy Log File used to capture the user access data i.e. it capture the pages that are being accessed by the user. Proxy server is in many-many cardinality since there are many users accessing many pages.

Server Log Files are in relationship of many to one since there is only one web server response to many users. Server log file do not record the cached pages requested.

Different types of Server Log File include:

- a. Referrer Log
- b. Error Log
- c. Agent Log
- d. Access Log

Referrer Log file contains information about the pages that is being referrer. **Error LogFile** records the errors of web site especially page not found error (404 File not found).

Agent Log File records the information about the website user's browser, browser version & Operating System.

Access Log file records all the click, hits and accesses made by the user to the website.

Web log File describes about different types of Log that can be captured. The Log captured should be stored in a specific format separated by comma, so that it can further use for processing. The Format which is used in the paper is Common Web log format [Table1] in a customized manner [2].

Table1 Attributes of log file and description

Attributes	Description
ClientIP	Client Machine IP address
Time	Time of Transaction
Date	Date When user made access
Server Site name	Internet Service name
Server Computer name	Server Name
Server Port	Server port configured for data transmission
Server Client Status	Status code returned by server 200, 404
User Agent	Browser Types that client used
Referrer	Link from where Client jump to site
Client Server URI Stem	Targeted Default Web page
URI Query	Client Query Which start "?"
Client Server Host	Host Header Name

\$ClientIP	, \$time,	\$scriptfilename,	\$serverport,
\$servername,	\$referrer,	\$statuscode,	\$uri,
\$useragent,	\$querystring,	\$host,	\$documentroot

Table2 Example

127.0.0.1, 2012-08-8 14:18:11, des/index.php, 80, local host, http://localhost/des/index.php?subj=1, 200, Mozilla/5.0 (X11; Linux x86_64) Chrome, subj=1, localhost, /var/www.

While we are capturing the log each attributes is separated by comma which make it easier retrieve of the attributes during preprocessing and the complete attributes are stored in the text file (.txt) which also allow for fastest retrieve of information from the server.

II. Data Preprocessing

Preprocessing converts the captured log data in to the valuable information which can be given for further pattern discovery. In this phase main steps includes are [1]:

1. Extraction the attributes from the web log which is located in web server
2. Cleaning the web logs and removing the redundant and irrelevant information.
3. Manage the data and put it in relational database or data warehouse.

2.1 Field Extraction

In Field Extraction we extract the attributes from the log file which is separated by character ‘.’. The attributes that are extracted is stored in the log table which is relational database. Process flow is shown in Fig.2.

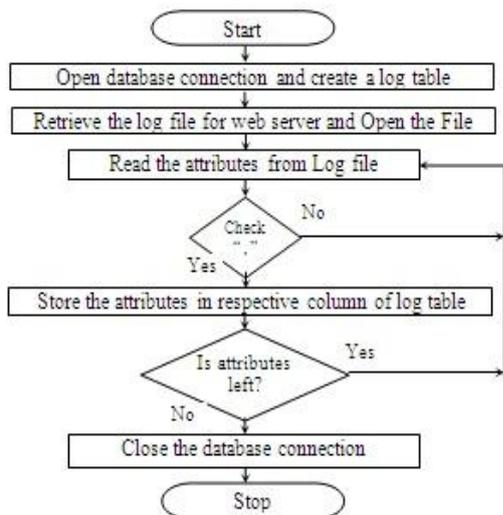


Fig.2 Process Flow for Field Extraction

2.2 Data Cleaning and Data Filtering

Data cleaning eliminates irrelevant or unnecessary records stored in log table. Since website will be accessed by millions of users. Data Cleaning is used to remove the records while analyzing the data. It removes all the jpg, gif and css files which the user has accessed & also the failed HTTP status code (404 Page Not Found). Data Filtering provides with the more better cleaning by removing the repeated pages that the user have accessed in the referrer string and also filtering help to reduce the path's accesses of

the page by splitting the Referrer String and getting valuable information.

By performing Data Cleaning and Filtering errors files, inconsistent data, missing data, repeatable data will be detected and removed to improve the quality of data. Process flow is shown in Fig.3.

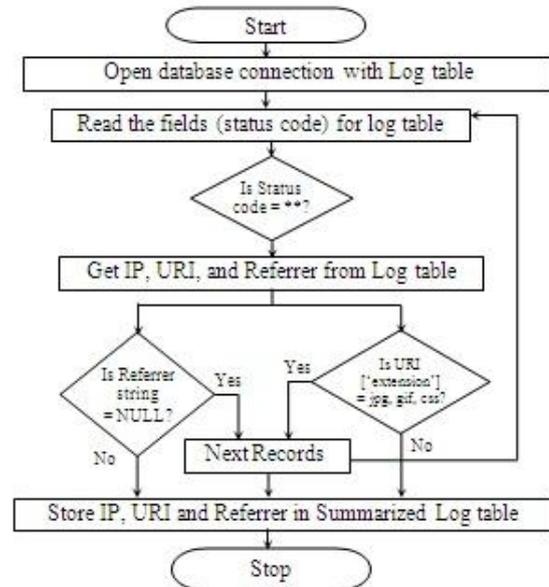


Fig.3 Process Flow of Data Cleaning and Filtering

2.3 User Identification

User Identification is based identify the user with the User Id which help to predict the user behavior uniquely i.e. pages the user have accesses along with the IP address of the user through which user was accessing the sites.

User Id of the user also helps in Session Identification i.e. if the session is set then the user is registered user and can start recording the pages the user accessing and also provide the user with dynamism of the pages but if session is not set the user is allow to access the pages but do not record any of the log information of the user which in turn reduces the log information of the non-registered user.

2.4 Data Summarization

In this paper data summarization is describe by the graphical representation of the log that are being captured. Since, useful and consistent records remain in the Summarized log table which helps in further pattern discovery process.

Summarized log table when compared with the log table provides information such as Total no. of records, Total no. of registered User, Total no. of URL, Visited count of the registered user.

III. Results

3.1 Comparison

We have captured a log for a month '19-07-2012' to '22-08-2012' of distance education system and depending upon the log captured we had following analysis. [Table3] shows the comparison between log table and summarized log table. [Table4] shows the Number of user that have uniquely access the distance education system. [Table5]

shows summarized statistical report. The Following Figures are showing total no. of records in Log table and summarized log table [Fig. 4], Number of access of the unique users [Fig. 5], File Size in (KB) [Fig. 6].

Table3 Comparison between Log Table (LT) and Summarized Log Table (SLT)

Attributes	LT	SLT
File Size (KB)	2154	519
Number of Records	343254	12519

Table4 Number of Access of the Unique Users

Client IP address	No. of Access
172.16.11.101	34
172.16.11.102	40
172.16.11.103	19
172.16.11.104	26
172.16.11.105	53
172.16.11.106	13
...	...
...	...

Table5 Summary Statistical Report

Status	Records
Errors	1289
Css	868
Jpeg	2548
Hits	1524785

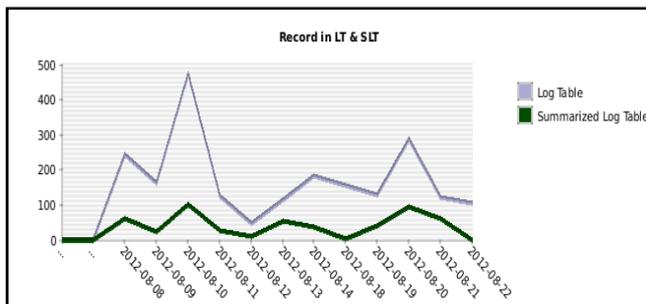


Fig.4 Total no. of records in Log table and Summarized log table (Date-wise)

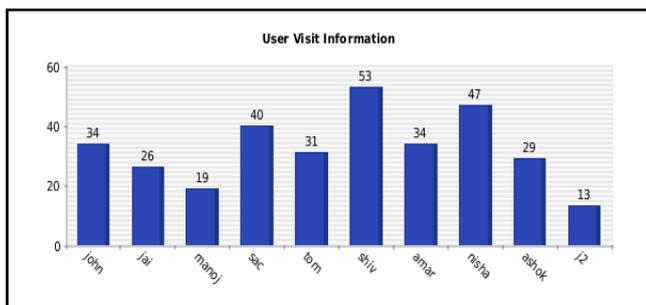


Fig.5 No. of accesses of unique users

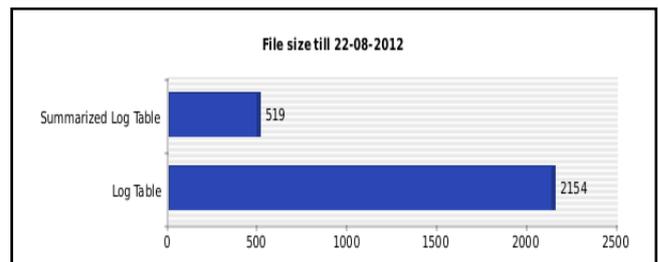


Fig.6 File size in (KB)

3.2 Analysis

From the Results after capturing the web log and representing it in graphical manner, we can get the conclusion as log file size is reduced to 1/4 of the actual log captured and record in the file are also correspondingly reduced.

IV. Conclusion

The information after Data Preprocessing can be given to pattern discovery process which includes different data mining techniques such as clustering, classification and association rules and since irrelevant information is removed therefore it speeds up the execution time and provides with valuable information to the users.

It also allows the user to have different access pattern due to which better accessibility is achieved. We can also record the information in the cookies and when the user ends its session we can then transfer the complete information from the cookies to the database.

REFERENCES

- [1] TheintT heintAye, Web log Cleaning for Mining of Web Usage Patterns, *IEEE 978-1- 61284-840-2/11*
- [2] Tasawar Hussian, Dr. Sohail Asghar, Dr. Hayyer Masood, Web Usage Mining: A Survey on Preprocessing of Web log File, *International Conference 978-1-4244080306/10*
- [3] Xinjin Li, Sujing Zhang, Application of Web usage mining in e-learning platform, *2010 International Conference on E-Business and E-Government*
- [4] R. Cooley, B. Mobasher, and J. Srivastava, Data preparation for mining World Wide Web browsing patterns, *Knowledge and Information Systems, Vol. 1, No. 1, 1999, pp. 5-32.*

Proceedings Papers:

- [5] Yuan, F., L.-J. Wang, et al. (2003). Study on Data Preprocessing Algorithm in Web Log Mining. Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2-5 November 2003

An Overview of Categorization techniques

B. Mahalakshmi¹, Dr. K. Duraiswamy²

¹Associate Prof., Dept. of Computer Science and Engineering, K. S. Rangasamy College of Technology, Tiruchengode, India

²Dean (Academic), K. S. Rangasamy College of Technology, Tiruchengode, India

Abstract : Categorization is the process in which ideas and objects are recognized, differentiated and understood. Categorization implies that objects are grouped into categories, usually for some specific purpose. A category illuminates a relationship between the subjects and objects of knowledge. The data categorization includes the categorization of text, image, object, voice etc. With the rapid development of the web, large numbers of electronic documents are available on the Internet. Text categorization becomes a key technology to deal with and organize large numbers of documents. Text representation is an important process to perform text categorization. A major problem of text representation is the high dimensionality of the feature space. The feature space with a large number of terms is not only unsuitable for neural networks but also easily to cause the over fitting problem. Text categorization is the assignment of natural language documents to one or more predefined categories based on their semantic content is an important component in many information organization and management tasks. This paper discusses various categorization techniques, tools and their applications in different fields.

Keywords- Clustering, Neural networks, Latent Semantic Indexing, Self-Organizing map.

I. Introduction

Automatic text categorization is an important application and research topic for the inception of digital documents. Text categorization [1] is a necessity due to the very large amount of text documents that humans have to deal with daily. A text categorization system can be used in indexing documents to assist information retrieval tasks as well as in classifying e-mails, memos or web pages in a yahoo-like manner.

The text classification task can be defined as assigning category labels to new documents based on the knowledge gained in a classification system at the training stage. In the training phase, given a set of documents with class labels attached and a classification system is built using a learning method, machine learning communities.

Text classification [4] tasks can be divided into two sorts: supervised document classification where some external mechanism provides information on the correct classification for documents, and unsupervised document classification, where the classification must be done entirely without reference to external information. There is also a semi-supervised document classification, where some documents are labeled by the external mechanism.

Text categorization [2] is the problem of automatically assigning predefined categories to free text documents. While more and more textual information is available online, effective retrieval is difficult without indexing and summarization of document content.

Document categorization is one solution to this problem. A growing number of statistical classification methods and machine learning techniques has been applied to text categorization including Neural Networks, Naïve Bayes classifier approaches, Decision Tree, Nearest neighbor classification, Latent semantic indexing, Support vector machines, Concept Mining, Rough set based classifier, Soft set based classifier[3].

Document classification techniques include:

- Back propagation Neural Network
- Latent semantic indexing
- Support vector machines
- Decision trees
- Naive Bayes classifier
- Self-Organizing Map
- Genetic Algorithm.

II. Back Propagation Neural Network

The back-propagation neural network [5] is used for training multi-layer feed-forward neural networks with non-linear units. This method is designed to minimize the total error of the output computed by the network. In such a network, there is an input layer, an output layer, with one or more hidden layers in between them. During training, an input pattern is given to the input layer of the network. Based on the given input pattern, the network will compute the output in the output layer. This network output is then compared with the desired output pattern. The aim of the back-propagation learning rule is to define a method of adjusting the weights of the networks. Then, the network will give the output that matches the desired output pattern given any input pattern in the training set [7].

The training of a network by back-propagation involves three stages: the feed forward of the input training pattern, the calculation and back-propagation of the associated error, the adjustment of the weight and the biases. The main defects of the BPNN can be described as: slow convergence, difficulty in escaping from local minima, easily entrapped in network paralyse, uncertain network structure. In order to overcome the demerits of BPNN some techniques are introduced and it is mentioned below.

Cheng Hua Li and Soon Cheol Park introduced a new method called MRBP. MRBP (Morbidly neuron Rectify Back-Propagation neural network) [5]. This method is used to detect and rectify the morbidly neurons. This reformative BPNN divides the whole learning process into many learning phases. It evaluates the learning mode used in the phase evaluation after every learning phase. This can improve the ability of the neural network, making it more adaptive and robust, so that the network can more easily escape from a local minimum, and be able to train itself more effectively.

Wei Wang and Bo Yu proposed a combined method called MBPNN and LSA. The MBPNN [6] accelerates the training speed of BPNN and improve the categorization accuracy. LSA can overcome the problems caused by using statistically derived conceptual indices instead of individual words. It constructs a conceptual vector space in which each term or document is represented as a vector in the space. It not only greatly reduces the dimension but also discovers the important associative relationship between terms. The two methods to improve the speed of training for BPNN in order to improve the back propagation algorithm in terms of faster convergence and global search capabilities are:

- Introduce momentum into the network.
Convergence is sometimes faster if a momentum term is added to the weight update formulas
- Using adaptive learning rate to adjust the learning rate.
The role of the adaptive learning rate is to allow each weight to have its own learning rate, and to let the learning rate vary with time as training progress.

Latent semantic analysis (LSA) uses singular value decomposition (SVD) [8] technique to decompose a large term-document matrix into a set of k orthogonal factors, it can transform the original textual data to a smaller semantic space by taking advantage of some of the implicit higher-order structure in associations of words with text objects. These derived indexing dimensions, rather than individual words, can greatly reduce the dimensionality and have the semantic relationship between terms. So even two documents don't have any common words, we also can find the associative relationship between them, because the similar contexts in the documents will have similar vectors in the semantic space. The SVD used for noise reduction to improve the computational efficiency in text categorization and also LSA expanded term by document matrix used in conjunction with background knowledge in text categorization. The supervised LSA had been proposed to improve the performance in text categorization.

MBPNN overcomes the slow training speed problem in the traditional BPNN and can escape from the local minimum. MBPNN enhances the performance of text categorization. The introducing of LSA not only reduces the dimension, further improves its accuracy and efficiency. Bo Yu *et al.* [40] have proposed text categorization models using back-propagation neural network (BPNN) and modified back-propagation (MBPNN). A major problem of text representation is the high dimensionality of feature space. Dimensionality reduction and semantic vector space generation was achieved using a technique Latent Semantic Analysis (LSA). They have tested their categorization models using LSA on newsgroup dataset. They found that computation time for neural network with LSA method was faster than the neural network with VSM model. Further, the categorization performance of neural network using LSA was better than using VSM.

III. Latent Semantic Indexing

Latent Semantic Indexing (LSI) is an indexing and retrieval method that uses a mathematical technique called Singular Value Decomposition (SVD) to identify patterns in the relationships between the terms and concepts contained in an unstructured collection of text. LSI is based

on the principle that words that are used in the same contexts tend to have similar meanings. A key feature of LSI is its ability to extract the conceptual content of a body of text by establishing associations between those terms that occur in similar contexts.

LSI overcomes two of the most severe constraints of Boolean keyword queries: multiple words that have similar meanings (synonymy) and words that have more than one meaning (polysemy). Synonymy and polysemy are often the cause of mismatches in the vocabulary used by the authors of documents and the users of information retrieval systems. [8] As a result, Boolean keyword queries often return irrelevant results and miss information that is relevant.

LSI is also used to perform automated document categorization. In fact, several experiments have demonstrated that there are a number of correlations between the way LSI and humans process and categorize text [9]. Document categorization is the assignment of documents to one or more predefined categories based on their similarity to the conceptual content of the categories [8]. LSI uses example documents to establish the conceptual basis for each category. During categorization processing, the concepts contained in the documents being categorized are compared to the concepts contained in the example items, and a category is assigned to the documents based on the similarities between the concepts they contain and the concepts that are contained in the example documents. Dynamic clustering based on the conceptual content of documents can also be accomplished using LSI. Clustering is a way to group documents based on their conceptual similarity to each other without using example documents to establish the conceptual basis for each cluster. This is very useful when dealing with an unknown collection of unstructured text.

Yan Huang described about Text Categorization via Support Vector Machines (SVMs) approach based on Latent Semantic Indexing (LSI) [10]. Latent Semantic Indexing is a method for selecting informative subspaces of feature spaces with the goal of obtaining a compact representation of document. Support Vector Machines [3] are powerful machine learning systems, which combine remarkable performance with an elegant theoretical framework. The SVMs well fits the Text Categorization task due to the special properties of text itself. The LSI+SVMs frame improves clustering performance by focusing attention of Support Vector Machines onto informative subspaces of the feature spaces. LSI is an effective coding scheme and It captures the underlying content of document in semantic sense. SVMs well fit for text categorization task due to the properties of text. LSI+SVMs shows to be a promising scheme for TC task.

Chung-Hong Lee *et al.* described that an LSI is a technique for Information Retrieval, especially in dealing with polysemy and synonymy [11]. LSI use SVD process to decompose the original term-document matrix into a lower dimension triplet. The triple is the approximation to original matrix and can capture the latent semantic relation between terms. A novel method for multilingual text categorization using Latent Semantic Indexing is mentioned here. The centroid of each class has been calculated in the decomposed SVD space. The similarity threshold of categorization is predefined for each centroid. Test

documents with similarity measurement larger than the threshold will be labeled Positive or else would be labeled Negative. Experimental result indicated that the performance on the precision, recall is quite good using LSI technique to categorize the multi-language text.

Sarah Zelikovitz and Finella Marquez presented a work that evaluates background knowledge created via web searches might be less suitable. For some text classification tasks, unlabeled examples might not be the best form of background knowledge for use in improving accuracy for text classification using Latent Semantic Indexing (LSI) [12]. LSI's singular value decomposition process can be performed on a combination of training data and background knowledge. The closer the background knowledge is to the classification task, the more helpful it will be in terms of creating a reduced space that will be effective in performing classification. Using a variety of data sets, evaluate sets of background knowledge in terms of how close they are to training data, and in terms of how much they improve classification.

Antony Lukas *et al.* made a survey about document categorization using Latent semantic indexing [13]. The purpose of this research is to develop systems that can reliably categorize documents using the Latent Semantic Indexing (LSI) technology [11]. Categorization systems based on the LSI technology do not rely on auxiliary structures and are independent of the native language being categorized. Three factors led us to undertake an assessment of LSI for categorization applications. First, LSI has been shown to provide superior performance to other information retrieval techniques in a number of controlled tests [8]. Second, a number of experiments have demonstrated a remarkable similarity between LSI and the fundamental aspects of the human processing of language. Third, LSI is immune to the nuances of the language being categorized, thereby facilitating the rapid construction of multilingual categorization systems. The emergence of the World Wide Web has led to a tremendous growth in the volume of text documents available to the open source community. It had led to an equally explosive interest in accurate methods to filter, categorize and retrieve information relevant to the end consumer. Of special emphasis in such systems is the need to reduce the burden on the end consumer and minimize the system administration of the system. The implementation of two successfully deployed systems employing the LSI technology for information filtering and document categorization was described. The systems utilize in-house developed tools for constructing and publishing LSI categorization spaces.

Two-stage feature selection algorithm [32] is based on a kind of feature selection method and latent semantic indexing. Feature selection is carried out in two main steps. First, a new reductive feature space is constructed by a traditional feature selection method. In the first stage, the original features dimension is decreased from m to t . Second, features are selected by LSI method on the basis of the new reductive feature space that was constructed in the first stage. In the second stage, the features dimension is decreased from t to k . The feature-based method and semantic method are combined to reduce the vector space. The algorithm not only reduces the number of dimensions drastically, but also overcomes the

problems existing in the vector space model used for text representation.

I.Kuralenok and I. Nekrest'yanov [41] have considered the problem of classifying the set of documents into given topics. They have proposed a classification method based on the use of LSA to reveal semantic dependencies between terms. The method used the revealed relationships to specify the function of the topical proximity of terms, which was then used to estimate the topical proximity of documents. The results indicated a high quality of classification. The computation cost of this method was high at the initial stage and relatively cheap at the classification stage. However, considering the problem of clusterization of documents, unlike the classification problem, the topics of the groups are not given in advance.

IV. Support vector machines

SVMs are a set of related supervised learning methods used for classification and regression. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that predicts whether a new example falls into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. A support vector machine constructs a hyper plane or set of hyper planes in a high or infinite dimensional space, which can be used for classification, regression or other tasks. A good separation is achieved by the hyper plane that has the largest distance to the nearest training data points of any class since in general the larger the margin the lower the generalization error of the classifier [14].

Lukui Shi *et al.* proposed an algorithm combined nonlinear dimensionality reduction techniques with support vector machines for text classification. To classify documents, the similarity between two text documents is considered in many algorithms of text categorization. Here geodesic distance is used to represent the similarity between two documents. In this algorithm, high-dimensional text data are mapped into a low-dimensional space with the ISOMAP algorithm after geodesic distances among all documents are computed at first. Then the low-dimensional data are classified with a multi-class classifier based single-class SVM [15].

ISOMAP is a nonlinear dimensionality reduction technique, which generalizes MDS by replacing Euclidean distances with an approximation of the geodesic distances on the manifold. The algorithm is to compute the geodesic distances between points, which represent the shortest paths along the curved surface of the manifold. For neighboring points; the input space distance gives a good approximation to the geodesic distance. For objects, the geodesic distances can be approximated by a sequence of short hops between neighboring points.

The multi-class classifier based on single-class SVM can effectively treat multi-class classification problems. The efficiency of the classifier will be rapidly degraded when the dimension of data becomes greatly high. Usually, the dimension of text data is huge. To fast classify high-dimensional text data, it is necessary to decrease the

dimension of high-dimensional data before classifying text documents. It is a good selection to combine the above multi-class classifier with ISOMAP.

Montanes E *et.al.* described a wrapper approach with support vector machines for text categorization [16]. Text Categorization is the assignment of predefined categories to documents plays an important role in a wide variety of information organization and management tasks of Information Retrieval (IR). It involves the management of a lot of information, but some of them could be noisy or irrelevant and hence, a previous feature reduction could improve the performance of the classification. Here they proposed a wrapper approach. This approach is time-consuming and also infeasible. But this wrapper explores a reduced number of feature subsets and also it uses Support Vector Machines (SVM) [18] as the evaluation system; and these two properties make the wrapper fast enough to deal with large number of features present in text domains.

István Pilászy [17] gave a short introduction of text categorization (TC), and important tasks of a text categorization system. He also focused on Support Vector Machines (SVMs), the most popular machine learning algorithm used for TC.

Support Vector Machines (SVMs) have been proven as one of the most powerful learning algorithms for text categorization. Support vector machines (SVMs) [19] are a set of related supervised learning methods used for classification and regression. In simple words, given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that predicts whether a new example falls into one category or the other. Intuitively, an SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on [20]. Redundant features and high dimension are well-handled.

Linear Support Vector Machines (SVMs) [33] have been used successfully to classify text documents into set of concepts. The training time was taken with respect to each category by SVMlight, PSVM, SVMlin, and SVMperf on two corpuses. The training times of all other algorithms were higher than SVM light on both corpuses. On reuters-21578, the training time of PSVM is the least, and on assumed, both SVMlin and PSVM achieve less training time when compared with other algorithms. The order of computational complexity of PSVM scales with respect to dimensionality of the corpus. The solution of FPSVM can also be obtained by solving system of simultaneous linear equations similar to PSVM. PSVM maintains almost constant training time irrespective of the penalty parameter and categories. The performance of PSVM can greatly be improved by using it along with advanced feature selection/extraction methods like word clustering, rough sets.

Wenqian Shanghan *et al.* [38] have designed a novel Gini index algorithm to reduce the high dimensionality of the feature space. They have constructed a new measure function of Gini index to fit text categorization. Improved Gini index algorithm was evaluated using three classifiers: SVM, kNN, fkNN. The

performance of new Gini index was compared with feature selection methods Inf Gain, CrossEntropy, CHI, and Weigh of Evid. The results showed that the performance of new method was best in some dataset and inferior in another dataset. As a whole, they concluded that their improved Gini index showed better categorization performance.

V. Decision Tree

Decision tree learning [25], used in data mining and machine learning, uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. More descriptive names for such tree models are classification trees or regression trees. In these tree structures, leaves represent classifications and branches represent conjunctions of features that lead to those classifications. A decision tree can be used to visually and explicitly represent decisions and decision making. In data mining, a decision tree describes data but not decisions rather the resulting classification tree can be an input for decision making [24].

The text categorization performance of purely inductive method is used [23]. Two inductive learning algorithms are: Bayesian classifier and other one is Decision tree. Both the algorithms studied about indexing the data for document retrieval and also extraction of data from the text sources.

The Bayes rule is to estimate the category assignment probabilities and then assign to a document those categories with high probabilities. The decision tree use the algorithm DT-min 10: to recursively subdivide the training examples into subsets based on the information gain metric [21].

Maria Zamfir Bleyberg and Arulkumar Elumalai introduced a rough set method. It is founded on the assumption that with every object of the universe we associate some information. Objects characterized by the same information are similar in view of the available information about them. The indiscernibility relation generated in this way is the mathematical basis of rough set theory. Any set of all indiscernible objects is called an elementary set, and forms a basic granule of knowledge about universe. Any union of some elementary sets is referred as crisp set, otherwise the set is rough. In the rough set theory, any vague concept is replaced by a pair of precise concepts: the lower and the upper approximation of the vague concept. The learning methods based on rough sets, can be used to support flexible, dynamic, and personalized information access and management in a wide variety of tasks.

VI. Naive Bayes Classifier

A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes theorem with strong independence assumptions. A naive Bayes classifier assumes that the presence or absence of a particular feature of a class is unrelated to the presence or absence of any other feature. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. One can work with the naive Bayes model without believing in Bayesian probability or using any Bayesian methods. In spite of their naive design and apparently over-simplified

assumptions, naive Bayes classifiers have worked quite well in many complex real-world situations. An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix [26].

Jantima Polpinij and Aditya Ghose solved an ambiguity problem of software errors because much of the requirements specification is written in a natural language format. It is hard to identify consistencies because this format is too ambiguous for specification purposes. [27] A method for handling requirement specification documents which have a similar content to each other through a hierarchical text classification. The method consists of two main processes of classification: heavy classification and light classification. The heavy classification is to classify based on probabilistic text classification (Naïve Bayes), while light classification is to handle elaborate specification requirement documents by using the Euclidean Distance. Slimming down the number of requirements specification through hierarchical text classification classifying may yield a specification which is easier to understand. That means this method is more effective for reducing and handling in the requirements specification.

Dino Isa *et al.* [42] have designed and evaluated a hybrid classification approach by integrating the naive Bayes classification and SOM utilizing the simplicity of the naive Bayes to vectorize raw text data based on probability values and the SOM to automatically cluster based on the previously vectorized data. Through the implementation of an enhanced naive Bayes classification method at the front-end for raw text data vectorization, in conjunction with a SOM at the back-end to determine the right cluster for the input documents, better generalization, lower training and classification time, and good classification accuracy was obtained. The drawback of this technique is the fact that the classifier will pick the highest probability category as the one to which the document is annotated too.

VII. Self-Organizing Map

The SOM [28] is an unsupervised-learning neural-network method that produces a similarity graph of input data. It consists of a finite set of models that approximate the open set of input data, and the models are associated with nodes (neurons) that are arranged as a regular, usually 2-D grid. The models are produced by a learning process that automatically orders them on the 2-D grid along with their mutual similarity.

Cheng Hua Li and Soon Choel Park described two kinds of neural networks for text categorization [30], multi-output perceptron learning (MOPL) and back-propagation neural network (BPNN), and then a novel algorithm using improved back-propagation neural network is proposed. This algorithm can overcome some shortcomings in traditional back-propagation neural network such as slow training speed and easy to enter into local minimum. The training time and the performance, and tested three methods are compared. The results showed that the proposed algorithm is able to achieve high categorization effectiveness as measured by the precision, recall and F-measure.

Richard Freeman *et al.* [35] have investigated the use of self-organizing maps for document clustering. They have presented a hierarchical and growing method using a series of one-dimensional maps. The documents were represented using vector-space model. Dynamically growing one-dimensional SOM were allocated hierarchically to organize the give set of documents. The hierarchical structured maps produced were visualized easily as a hierarchical tree. The results showed a more intuitive representation of a set of clustered documents.

Nikolaos and Stavros [36] have introduced LSISOM method, for automatic categorization of document collections. The method LSISOM obtained word category histograms from the SOM clustering of the Latent Semantic Indexing representation of document terms. The problem of high dimensionality of VSM word histograms document representation was suppressed by LSI representation. The SOM used was a two-dimensional SOM. They used 420 articles as dataset from the TIME Magazine. They have proved that LSISOM method is computationally efficient due to dimensionality reduction using LSI of documents. They have compared Standard SOM (SSOM) and LSISOM for a set of documents. They justified that consistent mapping of documents onto a single cluster was obtained by LSISOM.

The method topological organization of content (TOC) [37] is topology preservation of neural network for content management and knowledge discovery. TOC generate taxonomy of topics from a set of unstructured documents. TOC is a set of 1D-growing SOMs. The TOC method produced a useful hierarchy of topics that is automatically labeled and validated at each level. This approach used entropy-based BIC to determine optimum number of nodes. TOC and 2D-SOM were compared; the results showed that topological tree structure improved navigation and visualization. The main advantages of the approach are the validation measure, scalability, and topology representation. To improve TOC, feature selection method LSA can be used to enhance the association between terms.

Yan Yu *et al.* [39] have presented a new document clustering method based on one-dimensional SOM. This method obtained the clustering results by calculating the distances between every two adjacent MSPs (the most similar prototype to the input vector) of well trained 1D-SOM. Their work proved that procedure using 1D-SOM is simple and easy relative to that with 2D-SOM.

Tommy W. S. Chow and M. K. M. Rahman [43] have proposed a new document retrieval (DR) and plagiarism detection (PD) system using multilayer self-organizing map (MLSOM). Instead of relying on keywords/lines, the proposed scheme compared a full document as a query for performing retrieval and PD. The tree-structured representation hierarchically includes document features as document, pages, and paragraphs. MLSOM, a kind of extended SOM model, was developed for processing tree-structured data. A tree data consists of nodes at different levels. In MLSOM, there were as many SOM layers as the number of levels in the tree. They mapped the position vectors of child nodes into the SOM input vector. The mapping of position vectors was conducted using a simple 1D- SOM that is trained. Experimental results using MLSOM were compared against

tree-structured feature and flat-feature. They have shown that tree-structured representation enhanced the retrieval accuracy and MLSOM served as an efficient computational solution. However, for a very large scale implementation of DR and PD, it is difficult to process all documents in a single MLSOM module.

VIII. Genetic Algorithm

Genetic Algorithm is a search technique based on the principles of biological evolution, natural selection, and genetic recombination. They simulate the principle of 'survival of the fittest' in a population of potential solutions known as chromosomes. Each chromosome represents one possible solution to the problem or a rule in a classification.

The population evolves over time through a process of competition whereby the fitness of each chromosome is evaluated using a fitness function. During each generation, a new population of chromosomes is formed in two steps. First, the chromosomes in the current population are selected to reproduce on the basis of their relative fitness. Second, the selected chromosomes are recombined using idealized genetic operators, namely crossover and mutation, to form a new set of chromosomes that are to be evaluated as the new solution of the problem. GAs are conceptually simple but computationally powerful. They are used to solve a wide variety of problems, particularly in the areas of optimization and machine learning [29].

Clustering is an efficient way of reaching information from raw data and K-means is a basic method for it. Although it is easy to implement and understand, K-means has serious drawbacks. Hongwei Yang had presented an efficient method of combining the restricted filtering algorithm and the greedy global algorithm and used it as a means of improving user interaction with search outputs in information retrieval systems [31]. The experimental results suggested that the algorithm performs very well for Document clustering in web search engine system and can get better results for some practical programs than the ranked lists and k-means algorithm.

Wei Zhao *et.al.* introduced a new feature selection algorithm in text categorization [34]. Feature selection is an important step in text classification, which selects effective feature from the feature set in order to achieve the purpose of reduce feature space dimension. Genetic algorithm (GA) optimization features are used to implement global searching, and k-means algorithm to selection operation to control the scope of the search, which ensures the validity of each gene and the speed of convergence. Experimental results show that the combination of GA and k-means algorithm reduced the high feature dimension, and improved accuracy and efficiency for text classification.

IX. Conclusion

This paper discusses about various classification algorithms, their merits and demerits. The data categorization includes the categorization of text, image, object, voice etc. The focus of survey is done mainly on text categorization. The representation techniques, supervised and unsupervised classification algorithms and their applications are discussed. The survey has shown that different techniques exist for the problem. The research

should be still concentrated on efficient feature selection and on categorizing different types of data in different fields. In order to improve the text categorization various other semantic based machine learning algorithms can be added in future.

References

- [1] M.-L. Antonie and O. R. Za'iane, "Text document categorization by term association", In Proc. of the IEEE 2002 International Conference on Data Mining", pp.19–26, Maebashi City, Japan, 2002.
- [2] Yiming Yang and Jan O. Pedersen, "A Comparative Study on Feature Selection in Text Categorization", CiteSeerX, 1997.
- [3] Thorsten Joachims, "Text categorization with Support Vector Machines: Learning with many relevant features, Machine Learning: ECML-98, Vol.1398, pp.137-142, 1998.
- [4] Nidhi and Vishal Gupta, "Recent Trends in Text Classification Techniques", International Journal of Computer Applications, Vol.35, No.6, 2011.
- [5] Cheng Hua Li and Soon Cheol Park, "A Novel Algorithm for Text Categorization Using Improved Back-Propagation Neural Network", Springer, pp. 452 – 460, 2006.
- [6] Wei Wang and Bo Yu, "Text categorization based on combination of modified back propagation neural network and latent semantic analysis", *Neural Comput & Applic*, Springer Link, Vol. 18, No.8, pp.875–881, 2009.
- [7] Wei Wu, Guorui Feng, Zhengxue Li, and Yuesheng Xu, "Deterministic Convergence of an Online Gradient Method for BP Neural Networks", IEEE Transactions on Neural Networks, Vol.16, NO.3, 2005.
- [8] Christos H. Papadimitriou, Prabhakar Raghavan, Hisao Tamaki and Santosh Vempala, "Latent Semantic Indexing: A Probabilistic Analysis", Journal of Computer and System Sciences, Vol.61, pp.217_235, 2000.
- [9] S.T. Dumais, G. W. Furnas, T. K. Landauer, and S. Deerwester, "Using latent semantic analysis to improve information retrieval", Proceedings of CHI'88: Conference on Human Factors in Computing, ACM, pp.281_285, 1988.
- [10] Yan Huang, "Support Vector Machines for Text Categorization Based on Latent Semantic Indexing", Electrical and Computer Engineering Department, isn.ucsd.edu, 2003.
- [11] Chung-Hong Lee, Hsin-Chang Yang and Sheng-Min Ma, "A Novel Multilingual Text Categorization System using Latent Semantic Indexing", Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06), 2006.
- [12] Sarah Zelikovitz and Finella Marquez, "Evaluation of Background Knowledge for Latent Semantic Indexing Classification", American Association for Artificial Intelligence, 2005.
- [13] Anthony Zukas and Robert J. Price, "Document Categorization Using Latent Semantic Indexing", Symposium on Document Image Understanding Technologies, 2003.
- [14] Daniela Giorgetti and Fabrizio Sebastiani, "Multiclass Text Categorization for Automated Survey Coding", SAC 2003.

- [15] Lukui Shi, Jun Zhang, Enhai Liu, and Pilian He, "Text Classification Based on Nonlinear Dimensionality Reduction Techniques and Support Vector Machines", Third International Conference on Natural Computation, IEEE Xplore, Vol.1, pp.674-677, 2007.
- [16] Montanes E., Quevedo J. R. and Diaz I., "A Wrapper Approach with Support Vector Machines for Text Categorization", Springer, LNCS 2686, pp. 230-237, 2003.
- [17] István Pilászy, "Text Categorization and Support Vector Machines", Proceedings of the 6th International Symposium of Hungarian Researchers on Computational Intelligence, 2005.
- [18] Manabu Sassano, "Using Virtual Examples for Text Classification with Support Vector Machines", Journal of Natural Language Processing, Vol.13, No.3, pp. 21-35. 2006.
- [19] A. Basu, C. Watters and M. Shepherd, "Support Vector Machines for Text Categorization", Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03), Vol.4, pp.103.3, 2003.
- [20] Edda Leopold and Jorg Kindermann, "Text Categorization with Support Vector Machines: How to Represent Texts in Input Space?", Machine Learning, Vol.46, Nr.1-3, pp.423-444, 2002.
- [21] Chidanand Apt, Fred Damerou and Sholom M. Weiss, "Automated Learning of Decision Rules for Text Categorization", ACM Transactions on Information Systems, 1994.
- [22] Srinivasan Ramaswamy, "Multiclass Text classification A Decision Tree based SVM Approach", CS294 Practical Machine Learning Project, Citeseer, 2006.
- [23] David D.Lewis and Mark Ringuette, "A comparison of two learning algorithms for text Categorization, Symposium on Document Analysis and Information Retrieval", 1994.
- [24] C. Apte, F. Damerou, and S.M. Weiss, "Text Mining with Decision Trees and Decision Rules", Conference on Automated Learning and Discovery Carnegie-Mellon University, 1998.
- [25] Nerijus Remeikis, Ignas Skucas and Vida Melninkaite, "A Combined Neural Network and Decision Tree Approach for Text Categorization", Information Systems Development, Springer, pp.173-184, 2005.
- [26] P.Bhargavi and Dr.S.Jyothi, "Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils", IJCSNS International Journal of Computer Science and Network Security, VOL.9, No.8, 2009.
- [27] Mohamed Aly, "Survey on Multiclass Classification Methods", Neural Networks, 2005.
- [28] Teuvo Kohonen, Samuel Kaski, Krista Lagus, Jarkko Salojärvi, Jukka Honkela, Vesa Paatero, and Antti Saarela, "Self Organization of a Massive Document Collection", IEEE Transactions on Neural Networks, Vol.11, No. 3, pp.574-585, 2000.
- [29] Xiaoyue Wang, Zhen Hua and Rujiang Bai, "A Hybrid Text Classification model based on Rough Sets and Genetic Algorithms", SNPD '08. Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, IEEE Xplore, pp.971-977, 2008.
- [30] Cheng Hua Li and Soon Choel Park, "Text Categorization Based on Artificial Neural Networks", Neural Information Processing, Springer, Vol.4234, pp.302 – 311, 2006.
- [31] Hongwei Yang, "A Document Clustering Algorithm for Web Search Engine Retrieval System", International Conference on e-Education, e-Business, e-Management and e-Learning, IEEE, pp.383-386, 2010.
- [32] Jiana Meng and Hongfei Lin, "A Two-stage Feature Selection Method for Text Categorization", Seventh International Conference on Fuzzy Systems and Knowledge Discovery, IEEE, pp.1492-1496, 2010.
- [33] M. Arun Kumar and M. Gopal, "An Investigation on Linear SVM and its Variants for Text Categorization", Second International Conference on Machine Learning and Computing, IEEE, pp.27-31, 2010.
- [34] Wei Zhao, Yafei Wang and Dan Li, "A New Feature Selection Algorithm in Text Categorization", International Symposium on Computer, Communication, Control and Automation, IEEE, pp.146-149, 2010.
- [35] Richard Freeman, Hujun Yin and Nigel M. Allinson, "Self-Organizing Maps for Tree View Based Hierarchical Document Clustering", IEEE Xplore, pp.1906-1911, 2002.
- [36] Nikolas Ampazis and Stavros J. Perantonis, "LSISOM - A Latent Semantic Indexing Approach to Self-Organizing Maps of Document Collections", Neural Processing Letters 19, pp. 157-173, 2004.
- [37] Freeman R.T. and Hujun Yin, "Web Content Management by Self Organization", IEEE transactions on Neural Networks, Vol.16, No.5, pp.1256-1268, 2005.
- [38] Wenqian Shang, Houkuan Huang, Haibin Zhu, Yongmin Lin, Youli Qu and Zhihai Wang, "A novel feature selection algorithm for text categorization", Expert Systems with Applications 33, pp.1-5, 2007.
- [39] Yan Yu, Pilian He, Yushan Bai and Zhenlei Yang, "A Document Clustering Method Based on One-Dimensional SOM", Seventh IEEE/ACIS International Conference on Computer and Information Science, pp.295-300, 2008.
- [40] BoYu, Zong-ben Xu and Cheng-hua Li, "Latent semantic analysis for text categorization using neural network", Knowledge-Based Systems, Vol.21, pp.900-904, 2008.
- [41] Kuralenok I. and Nekrest'yanov I., "Automatic Document Classification Based on Latent Semantic Analysis", Programming and Computer Software, Vo.26, No.4, pp.199-206, 2000.
- [42] Dino Isa, Kallimani V.P. and Lam Hong Lee, "Using the self organizing map for clustering of text documents", Expert Systems with Applications, Vol.36, pp.9584-9591, 2009.
- [43] Tommy W.S. Chow and Rahman M.K.M., "Multilayer SOM with Tree-Structured Data for Efficient Document Retrieval and Plagiarism Detection", IEEE Transactions On Neural Networks, Vol.20, No.9, pp.1385-1402,2009.

Review of Advanced EGR and Breathing Systems for High Performance and Low Emission HSDI Diesel Engine

Pundlik R Ghodke¹, Dr. Jiwak G. Suryavanshi²

Research Scholar, Mechanical Engineering Department, VNIT Nagpur / Dy. General Manager, Mahindra Research Valley, Chennai, India

Associate Professor, Mechanical Engineering Department, VNIT, Nagpur, India

ABSTRACT: In The Future, Beyond 2010 Additional Measures Is Required To Be Taken In To Account For Careful Selection Of Egr And Boosting Technology In Conjunction With After-Treatment System. At The End The Cost To Benefit Ratio Of After-Treatment System And The Performance Of Egr-Boost Systems Will Determine The Future Diesel Engine System Configurations. Diesel Particulate Filter (Dpf) Became As Essential Element Herein Beyond Euro 4 Emissions. This Paper Deals With Different Engine Boosting Systems Taking Into Account The Special Needs Of Turbochargers And The Interaction Between Egr And Boosting Technology In Conjunction With Selection Of After-Treatment System. The Investigation Has Been Done With A Calibrated Simulation Model For Selection Based On Benefit Verses Cost Of The System.

Key words: DPF; HP EGR; LP EGR; NO_x; PM; VGT Turbocharger

I. INTRODUCTION

All modern diesel engines with passenger car or truck application, Selection of EGR, boosting system and after-treatment systems together is key for engine performance and cost benefit to meet future stringent emission norms beyond Euro4.

Figure 1 shows various options of after-treatment systems in conjunction with EGR. Layout of EGR affects the engine out NO_x and PM emissions and performance of turbocharger due to change in exhaust flow rate going over the turbine. Also back pressure due to after-treatment system influence the turbocharger efficiency and air flow and EGR rate. This effect the engine performance, fuel economy and engine out emissions. Hence lot of attention is required while selection of EGR layout, intake system layout and exhaust system layout.

This was done with building simulation model and assessing the performance with various combinations before design stage.

	Reduction potential	HC	CO ₂	Costs	Risks
EURO4 EGR → DOC → DPF					
High-EGR combustion with DPF	⊙	⊙	⊙	⊙	HC Emissions Coking / carbonisation Combustion noise
High-EGR combustion with DPF	⊙	⊙	⊙	⊙	Development effort BSFC penalty
UREA-SCR System EGR → DOC → DPF → SCR	⊙	⊙	⊙	⊙	UREA consumption Costs / package
Lean Nox Trap System EGR → NSC → DPF	⊙	⊙	⊙	⊙	Aging / durability Regeneration strategy Costs

Fig1: Diesel Engine with different EGR and after-treatment system Layout.

II. EGR-BOOST CONCEPT

High pressure EGR (HP EGR) take the exhaust before the turbine via an ECU controlled EGR-valve to the air intake as shown in figure 2. For better breathing efficiency, EGR-coolers are used. EGR cooler reduces EGR gas temperature and it helps increase in volumetric efficiency of engine. With addition of exhaust gas oxygen content reduced and it help reduction of NO_x formation at part load condition. In addition, the application of gas/water charge air coolers may meet special package requirements. Niche applications today require even 2 stage cooling circuits to achieve the best thermodynamic results as shown in figure 5. Deposit formation is an issue which cannot be totally prevented. For that reason the untreated exhaust should not pass the charge air cooler which excludes its usage during part load conditions.

Classical HP-EGR reduces the turbine gas flow. The energy, driving the compresses or, decreases while the necessary compressor pressure ratio to keep the engine running at the same load points. Closing the Variable turbine geometry (VGT) increases the exhaust back pressure and the energy provided to the turbocharger. Mild EGR concepts used for Euro 4 application this kind of strategy is absolutely elegant and sufficient.

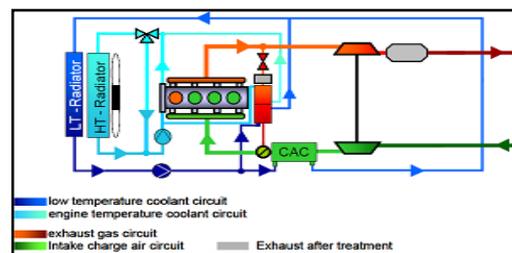


Fig 2: High Pressure EGR-Boost System

However, as more EGR rate is required, the turbocharger cannot be keep up and the operation point in the compressor map moves toward the surge line as shown in figure 3. For advanced EGR concepts where more stringent emission of Euro 5 , the turbocharger may not be able to deliver the necessary mixture of fresh air and exhaust to the engine. There are different matching opportunities to improve this situation, but engine power output and fuel efficiency is going to suffer.

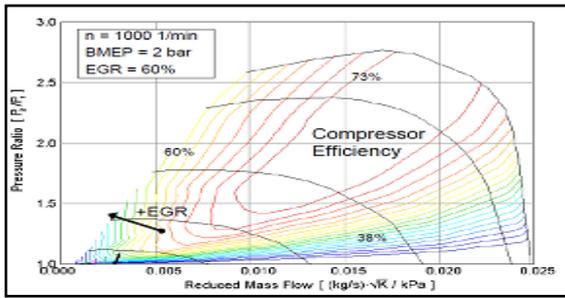


Fig 3: Influence of High Pressure EGR on compressor map.

2.1 Low Pressure Egr System

Low pressure loop EGR system is a well known technology that offers an alternative to meet the requirements mentioned above. The exhaust is taken after turbine and introduced in front of the compressor as shown in figure 4. Before the market introduction of Diesel particulate filter (DPF), the entire air intake system including the compressor would have been subject to deposit formation. This was one of the main reasons that limited the application of this kind of system.

The DPF can be therefore seen as enabler of LP-EGR. The exhaust is cooler as compare to HP-EGR is clean. It is expected, even after recompression, that the thermal behavior is superior to HP-EGR as the charge air cooler (CAC) is used in addition.

The LP-loop, like the HP-loop, is equipped with an EGR Valve. To increase the pressure difference between exhaust and air intake, especially to drive higher rates, an exhaust throttle is needed which will be closed to increase exhaust back pressure thereby increasing EGR.

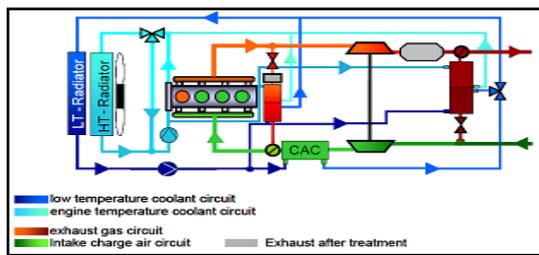


Fig 4: Low Pressure EGR system layout

2.2 Thermodynamic Comparison Of Lp And Hp Egr System

Figure 5 shows the introduction of the LP EGR-path. The engine has been operated at 2500 RPM at a load of 12 bar BMEP and an EGR rate of 30%. The very left line in the map represents the HP-mode. The other vertical lines represent increasing amounts of LP Loop EGR. For each of the lines, VTG-position is varied from open (Bottom) to closed (TOP). The very right, nearly vertical lines expresses the same as before but under LP-mode. The lines in between are splits of HP-and LP-loop systems. The overall EGR –rate does not change. It can be clearly seen that the air mass flow is increased when replacing HP-EGR by LP-EGR. The turbine flow is also influenced positively and better compressor efficiencies are expected obviously.

EGR-cooling helps to decrease the necessary pressure ratio and allows opening the VTG leading to better turbine efficiencies.

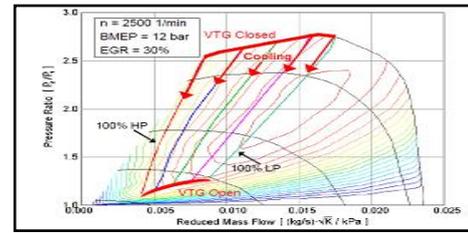


Fig 5: Influence of VGT, EGR-split and EGR-cooling

The engine parameters like air/ fuel ratio, BMEP and BSFC can be analyzed in figure 6, 7 and 8. Areas of air/fuel ratio smaller than 1 have been excluded. The results are based on the relationship of turbine efficiency depending on the VTG-position, pumping losses as part of boost pressure, the efficiency chain turbine-compressor and the necessity to add energy by fuel, represented by lambda. Closed VTG and a high turbine flow (LP-EGR) increase the pumping losses in principal. The corresponding lambda is high due to availability of excess air in the combustion chamber. At the end this leads to the best BSFC at the air fuel ratio of one for the widest open possible VTG position, represented by the right bottom corner of the operation maps within the compressor mappings. The best distance from the smoke limit will be reached by the higher boost pressures created by closing VTG. This has to be paid by a slightly higher BSFC caused by higher pumping losses.

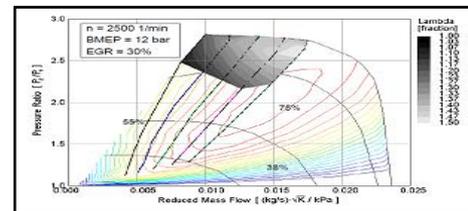


Fig 6 : Air Excess Ratio at 2500 rpm, 12 bar BMEP and 30% EGR

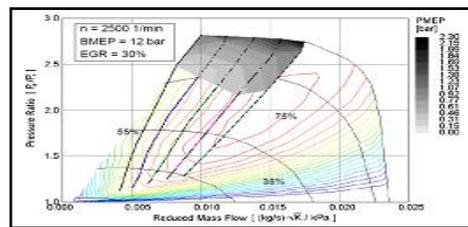


Fig 7: Pumping Losses @ 2500 rpm, 12 bar BMEP and 30 % EGR rate

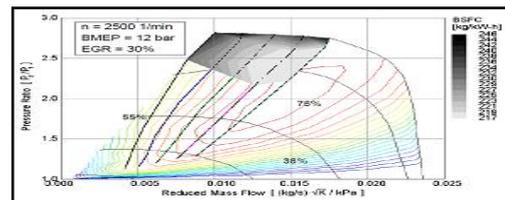


Fig 8: Specific Fuel consumption @ 2500 rpm, 12bar BMEP and 30 % EGR rate

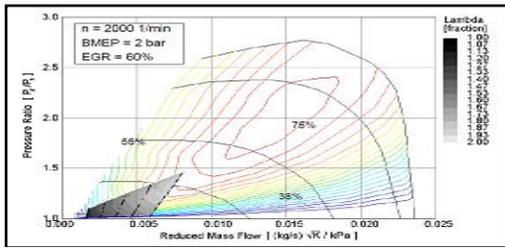


Fig 9: Air Excess Ratio @ 2000 rpm, 2 bar BMEP and 60% EGR rate

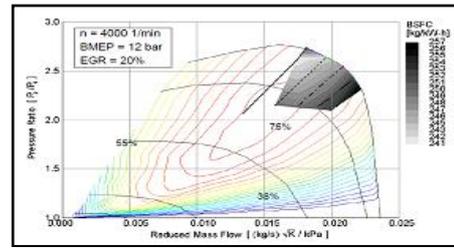


Fig 14: Specific Fuel Consumption @ 4000 rpm and 20% EGR rate

For low load conditions @ 2000 rpm and 2 bar BMEP and high EGR rate of 60 % same results are seen. Refer figure 9, 10 and 11. In addition, the benefit of higher exhaust gas flow can be easier seen even where the turbocharger does not build up a significant boost pressure. The significance between boost pressure and lambda is higher as both turbine and compressor are just starting to work.

Two operations are mentioned above are important for pass car diesel engines due to the requirements in the emission test procedure. Truck applications have to deal with EGR at high speeds and loads. As the compressor efficiencies are decreasing by higher flows and turbine conditions aren't improved, either the best BSFC are reached on the HP-loop mode area. An EGR-split of 25% LP and 75% HP flow seems to be the best operation point. Figure 6 to 14 shows an advantage of a split between HP and LP EGR. The split just optimizes the exhaust gas flow as a best fit to the turbine characteristics. At low engine speeds/loads LP helps to increase the energy flow to the turbine. At high speeds HP EGR may help to avoid turbine efficiency deterioration by wide VTG position.

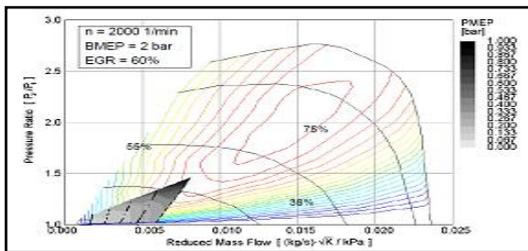


Fig 10: Pumping Losses @ 2000 rpm, 2 bar BMEP and 60% EGR Rate

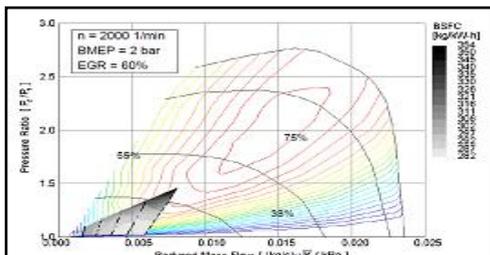


Fig 11: Specific Fuel Consumption @ 2000 rpm and 60% EGR rate

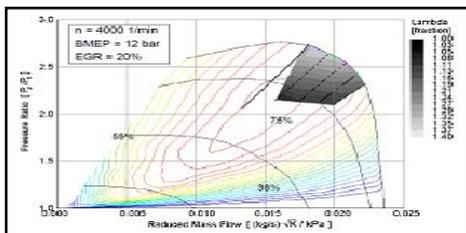


Fig 12 : Air Excess Ratio @ 4000 rpm and 20% EGR Rate

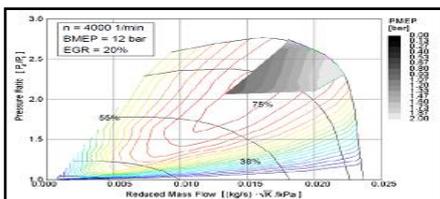


Fig 13 : Pumping Losses @ 4000 rpm, 12 bar BMEP and 20 % EGR rate

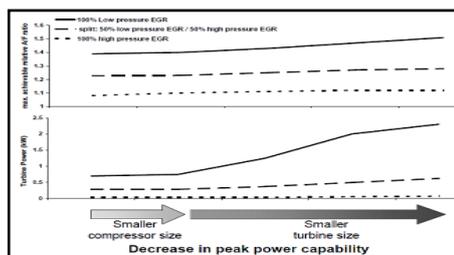


Figure 15: EGR-Split Turbo matching, @ 1500 rpm, 2 bar BMEP and 60% EGR rate

IV. TRANSIENT BEHAVIOUR

Low pressure EGR (LP-EGR) systems comprise a large air intake volume than high pressure EGR (HP-EGR) systems. Especially applications with under floor DPF will show a huge difference. Most of the discussions therefore focus on this possible disadvantage of LP-EGR systems

when it comes to the transient behavior of the diesel engine. The additional pipes and volumes have to be emptied of the air/exhaust of the previous operation point which might require some time.

For that reason the comparison was run to find the relationships between HP- and LP-system-acceleration times shown in figure 16. The load step goes from 2 bar BMEP to 9 bar, including a reduction in EGR. This load step represents part of the Euro 5 where especially small engines in heavy vehicles are facing the NOx-challenge. These 3-Charts in figure 16 present the BMEP, the necessary EGR-rate and the lambda. The graphs show the differences between LP and HP, including different control strategies for HP.

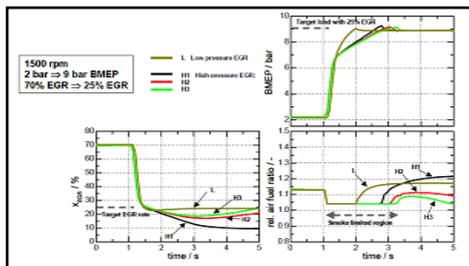


Fig 16: Load change @ 1500 rpm (H1: Strategy to achieve highest boost pressure, H2 compromise between H1 and H3 and H3: strategy to achieve highest EGR)

All the strategies reach the soot limit of lambda equal to 1.05 easily. The low LP -EGR system reaches the required BMEP first and shows a nearly stable EGR-rate. Strategy H1, with the focus on the fast increasing boost pressure, needs approximately one sec to catch up and does not reach the necessary EGR response but is running near the smoke limit: the operating point represents the borderline performance of a HP-system even at 9 bar BMEP. The example emphasizes the superior behavior of a LP-system at mid loads and low speeds considering future need of EGR-rates as mentioned before.

The diesel engine used has a power output of less than 45 kW/litre. The basic ability to handle EGR at low loads/speeds should be reasonable. Nevertheless LP shows significant advantages. Application with more than 50 kW/litre specific power output should therefore depend even more on the LP EGR-system properties.

The additional volume in the air intake of the LP EGR system can be seen as damping factor during transient during the filling process of a charged EGR-rate and boost pressure. Add on volume is created by the compressor and necessary tubes, which have been assumed to fill 5 to 6 liters. Figure 17 is extracts the BMEP curve of figure 16. The disadvantage caused by the LP-add on volume is negligible compared to the advantage in BMEP rise after 1.4 sec. The higher initial turbocharger speed, enabled by the higher exhaust mass flow through the turbine, ensures a smaller “perceivable” load step for the turbocharger.

The simplified explanation for the enhanced dynamic response is a sort of replacement of EGR by fresh air. In that defined load step the compressor is just considering rather small increase in turbocharger speed. This LP-Advantage in transient response corresponds to the controllability.

In addition, the interaction with VTG position is much less important than in a HP system. The result can be seen in figure 16 where the targeted EGR rate is reached and kept with high precision.

When EGR rates will increase in future emissions regulation the demand of more responsive LP-systems will increase as shown in figure 18. The higher the EGR rate, the larger the difference in response time based on the dominant energy balance at the turbine.

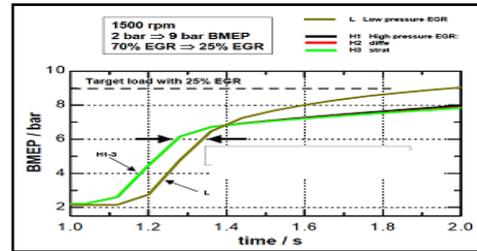


Figure 17: Load change @ 1500 rpm

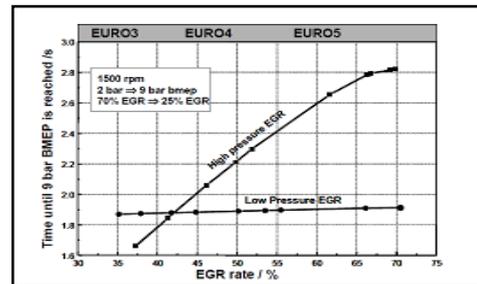


Fig 18 : Load Change for different EGR Rate and EGR-Split

V. CONCLUSION AND CONCEPT COMPARISON OF HP VERSES LP EGR

High pressure EGR	Low pressure EGR
<ul style="list-style-type: none"> Proven / developed system Good BSFC at low speeds/loads possible (low pumping losses) Intake throttling necessary for higher EGR-Rates Dynamics suffer due to low turbine speeds Low possible λ due to low intake air density and boost pressures deficit Full load EGR-rate limited by EGR-cooling and turbo charger capacity 	<ul style="list-style-type: none"> Higher EGR-rate at same λ in all map areas possible “clean” EGR (no soot, HC contamination) Near perfect EGR-distribution even at high EGR-rates (“HCCI” enable) High boost pressures with EGR possible control of LP-EGR fraction less coupled to turbo charger control Smaller necessary EGR-cooler / front radiator capacity through better use of charge-air cooler Higher breathing volume Measurement of LP-EGR fraction difficult (when necessary using HP-EGR in Addition) Pressure ratio limited by compressor inlet temperature Acid condensation in the compressor / intake area

Fig 19: Comparison of LP and HP EGR-Boost System

Configuration	Reduction-potential NOx / CO2	Dynamics	weight	packaging	Contamination	Costs
HP-EGR cooled VTG EGR - DOC - DOC						
HP-EGR highly cooled VTG EGR - DOC - DPF	*	⊖	⊖	⊖	⊖	⊖
HP-EGR 2-stage cooled VTG EGR - DOC - DPF	**	⊖	⊖	⊖	⊖	⊖
LP-EGR 1-stage cooled VTG EGR - DOC - DPF	**	⊖	*	⊖	*/⊖	⊖
LP-EGR 2-stage cooled VTG EGR - DOC - DPF	***	⊖	*	⊖	*/⊖	⊖
UREA-SCR System EGR - DOC - DPF - SCR	***	⊖	⊖	⊖	⊖	⊖

Figure 20: Comparison of Future Emission Concept with constant Peak power output.

Based on these results, EGR-systems will enable diesel engines to breathe even more exhaust than today. LP-EGR systems will allow lower NOx-emissions without any

after-treatment, which helps system costs lower, weight and complexity down refer figure 19.

Adding a 2nd cooling stage will lowering NOx by similar air intake temperatures. To optimize the best configuration depending on the situation in the vehicle is the task how compromise between the parameters mentioned in figure 19.

The single stage cooled LP-EGR system offers the most advantages based on the number of points and parameters.

The established HP-EGR system offers advantages like compact packaging, low compressor intake temperatures even at high power outputs, and low exposure of components to water and acid. The thermodynamic disadvantages have been explained and limit the HP-capabilities significantly.

The LP-EGR system, enabled by the DPF is new and seems to consume package volume, which can be smart designs like closed coupled filters and catalysts. Water and acid precipitation are subject to current development to limit or exclude their impact in the future.

EGR-systems will enable diesel engines to breathe even more exhaust than today LP-EGR systems will allow lower NOx-emissions without any after-treatment, which will help to keep system costs, weight and complexity down.

References

- [1] HEYWOOD, J.B., Internal combustion Engines Fundamentals, Mc Graw-Hill, Inc., 1988
- [2] WATSON, N. and JANOTA, M.S., Turbo charging the internal combustion engine-wiley-interscience, 1982
- [3] JUNMIN, WANG and et.al.: 2008-01-1198: Development of High Performance Diesel Engine Compliant with Euro V Norms, 2008 World Congress, Detroit, Michigan April 2008
- [4] Ramdasi SS and Etal : 2011-26-0033, SIAT2011, 'Design and Development of 3 Cylinder 75 kW/Litre High Power Density Engine for Passenger Car Application to Meet EIV /E V Norms
- [5] Borg-Wirger Turbo & emission systems Modern turbo charging systems for pass car/Future breathing system requirements for clean diesel engines.
- [6] Steven D. Arnold: 2004-01-1354: Turbocharging Technologies to Meet Critical performance Demands of Ultra low Emissions Diesel Engines, 19th cliff Garret Turbomachinery Award lecture, SAE international.
- [7] Sylvain Saulnier and ET al : 2004-01-0929: Computational study of diesel Engine down sizing using two stage Turbocharging, SP-1826, 2004 SAE World Congress, Michigan.
- [8] Rabih Omran and Et al : 2008-01-1732 : Optimizations of the In cylinder Air Filling For Emission Reduction in Diesel Engines, June 2008 SAE International, Powertrains, Fuel and Lubricants Congress, Shanghai, CHINA.
- [9] Tie Li and Et al : 2008-01-0647 : Effect of Exhaust Catalysts on Regulated and Unregulated Emissions from Low Temperature Diesel Combustion with High Rates of Cooled EGR, SP 2168, Incylinder NOX and Particulate Control, 2008.
- [10] G Avolio and Et al : 2007-24-0014: Effect of Highly Cooled EGR on Modern Diesel Engine Performance at low Temperature Combustion Condition, ICE 2007, Kapri Naples, Italy, SAE international.
- [11] Naoya Ishikawa and Et al : JSE 20077302:2007-01-1881: Study of effects on EGR Cooler on Performance on Combustion Properties of the Pre-mixed Compression Ignition Combustion By multi Cylinder DI Diesel Engine, Society of Automotive Engineers of Japan.
- [12] Robert D Chalgren and Et al : 2002-01-0076 : Controlled EGR Cooling System for Heavy Duty Diesel Application Using Vehicle Cooling System, SAE 2002 World Congress, Detroit, Michigan.
- [13] Raffael Schubinger and Et al: 2001-01-3497: Influence of EGR on Combustion and Emissions of Heavy duty DI Diesel Engines Equipped with Common Rail system, International Fall Fuels and Lubricants Meetings and Expositions, San Antonio, Texas Sept 200129.
- [14] Carlos Castano and Et al: 2007-26-019: Advantages in the EGR Cooler Performance by Using Internal Corrugated Tubes Technology, SIAT 2007, and SAE international.
- [15] Yohan Chi and Et al : 2002-01-0504 : Effect of VGT and Injection Parameters on Performance on HSDI Diesel Engine with Common Rail FIE System, 2002 World Congress, Detroit, Michigan April 2002, SP 1696, Diesel Fuel Injection and Spray 2007
- [16] Peter Lanzerath and Et al : 2010-01-1212 : Investigations of Chemical Aging of Diesel Oxidation Catalyst and Coated Diesel Particulate Filter, SAE International
- [17] Alexzandar Sappok and Et al : 2010-01-0811: Ash Effect on Diesel Particulate Filter Pressure Drop Sensitivity to Soot and Implication for Regeneration Frequency and DPF Control, SAE International.
- [18] Keld Johanesson and et al : 2010-01-0559 : NO2 Reduction, Passive and Active Soot Regeneration Performance of Pd Base Metal Coating on SiC filter, SAE International.
- [19] Jim Parks and Et al: 2007-01-3997; Characterisation of In Cylinder Techniques For Thermal Management of Diesel After treatment. Power Trains Fuels and Lubricants Meeting, Rosemont , Illinois Oct 2007
- [20] Dr. Olaf Weber and Et al, Future Breathing System Requirements for Clean Diesel Engines, 2005
- [21] Oldrich vitek and Et al , Comparison of different EGR Solutions, 2008-01-206, SAE International

Definitions, Acronyms, Abbreviations

- LP-EGR : Low pressure exhaust gas circulation
HP-EGR : High Pressure exhaust gas circulation
DPF : Diesel particulate filter
BMEP : Brake Mean Effective pressure
EGR : Exhaust Gas Recirculation
BSFC : Brake specific fuel consumption
PM: Particulate matter
HC: Hydrocarbon
NOx: Oxides of Nitrogen
VGT : Variable Turbine Geometry

Implementation of Order Statistic Filters on Digital Image and OCT Image: A Comparative Study

A. Stella¹, Dr. Bhushan Trivedi²

(Department of M.C.A, S.K.Patel Institute of Computer Studies, Gandhinagar – 23, Gujarat, India)
(Director, GLS Institute of Computer Technology (MCA), Ahmedabad-6, Gujarat, India.)

Abstract : Optical Coherence Tomography (OCT) technology introduces speckle, an insidious form of multiplicative noise which degrades the quality of OCT images. In this paper order statistics filters like Mean, Median, Min, Max and Alpha Trimmed Mean are implemented on digital image and OCT images. The images were tested with two different mask sizes of 5*5 and 7*7. The objective evaluation of both the types of images was performed using various image metrics like peak signal to noise ratio, root mean square error and image quality index. However, the existing algorithms proved to be effective for digital images and not for OCT images.

Keywords: Despeckling, Image Metrics, Optical Coherence Tomography, Order Statistics Filter, Speckle Noise.

I. INTRODUCTION

Diagnostic imaging has become a remarkable tool for medical diagnosis and disease prevention. The available technologies like Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) can provide non-invasive images of the human body. These techniques however, cannot generate high resolution images because of their physical limitations. To overcome this limitation, optical imaging has been being actively developed with the ultimate goal of high resolution ultrafast in-vivo imaging. One such technology is, Optical Coherence Tomography (OCT) a method for imaging the internal structure of biological tissue in vivo with micron resolution. OCT is based on the coherence properties of light. It enables the real-time, in situ visualization of tissue microstructure without the need to excise and process a specimen as in conventional biopsy and histopathology. One of its main limitations is the presence of speckle noise which obscures small and low-intensity features. "Coherence" in OCT technology introduces speckle, an insidious form of noise which degrades the quality of OCT images. Speckle arises as a natural consequence of the limited spatial-frequency bandwidth of the interference signals measured in optical coherence tomography (OCT) [1]. In images of highly scattering biological tissues, speckle has a dual role as a source of noise and as a carrier of information about tissue microstructure. There are different techniques developed to remove the speckle from the OCT image [2].

OCT images, as well as all other imaging modalities that involve a coherent light source, are affected by speckle noise. Speckle, arising from constructive and destructive interferences of the backscattered waves appears as a

random granular pattern [1] that significantly degrades image quality and complicates further image processing tasks, like image segmentation and edge detection.

II. MATHEMATICAL MODEL OF SPECKLE

Speckle is well modeled by a multiplicative noise. It is a random signal where the average amplitude increases with the overall signal intensity. It appears as bright specks in the lighter region of the image. It can be modeled as a pixel value multiplied by the random value. Speckle noise can be modeled as:

$$Y(x, y) = S(x, y).N(x, y) \quad (1)$$

where Y, S and N represent the noisy data, signal and speckle noise, respectively. In order to change the multiplicative nature of the noise to additive one, a logarithmic transformation is applied to the image data [3]. Taking logarithm of the both sides of equation (1), leads to:

$$f(x, y) = s(x, y) + e(x, y) \quad (2)$$

where f, s and e represent logarithms of the noisy data, signal and noise, respectively.

III. ORDER STATISTICS FILTER

Order Statistics filters are nonlinear spatial filters which are based on ordering the pixels contained in an image. Usually, sliding window technique [4,5] is employed to perform pixel-by-pixel operation in a filtering algorithm. The local statistics obtained from the neighbourhood of the center pixel gives a lot of information about its expected value. If the neighbourhood data are ordered (sorted), then ordered statistical information is obtained. If this order statistics vector is applied to a finite impulse response (FIR) filter, then the overall scheme becomes an order statistics (OS) filter [1, 6]. They are differentiated based on how they choose the values in the sorted list.

Minimum and Maximum Filter

The minimum filter selects the smallest value within the pixel values and maximum filter selects the largest value within of pixel values. This is accomplished by a procedure [7] which first finds the minimum and maximum intensity values of all the pixels within a windowed region around the pixel. If the intensity of the central pixel lies within the intensity range spread of its neighbors, it is passed on to the output image unchanged. However, if the central pixel intensity is greater than the maximum value, it is set equal to the maximum value; if the central pixel intensity is less than the minimum value, it is set equal to the minimum value. The minimum and maximum filters are represented as follows:

$$\hat{f}(x, y) = \min_{(s,t) \in S_{xy}} \{g(s, t)\} \quad (3)$$

$$\hat{f}(x, y) = \max_{(s,t) \in S_{xy}} \{g(s, t)\} \tag{4}$$

Median filter

The median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value [6, 7]. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.)

$$\hat{f}(x, y) = \text{median}\{g(s, t)\}_{(s,t) \in S_{xy}} \tag{5}$$

Alpha Trimmed Mean Filter

The alpha-trimmed mean (ATM) filter [4,7] is based on order statistics and varies between a median and mean filter. It is so named because, rather than averaging the entire data set, a few data points are removed (trimmed) and the remainders are averaged. The points which are removed are most extreme values, both low and high, with an equal number of points dropped at each end (symmetric trimming). In practice, the alpha-trimmed mean is computed by sorting the data low to high and summing the central part of the ordered array. The number of data values which are dropped from the average is controlled by trimming parameter alpha which is being expressed as:

$$\hat{f}(x, y) = \frac{1}{mn-d} \sum_{(s,t) \in S_{xy}} g_r(s, t) \tag{6}$$

IV. IMAGE METRICS

The quality of an image is examined by objective evaluation as well as subjective evaluation. For subjective evaluation, the image has to be observed by a human expert. The human visual system (HVS) [8] is so complicated that it is not yet modeled properly. There are various metrics used for objective evaluation of an image. Some of them are mean squared error (MSE), root mean squared error (RMSE), and peak signal to noise ratio (PSNR) [9]. The universal image quality index (IQI) [10] is modeled by considering three different factors: (i) loss of correlation, (ii) luminance distortion and (iii) contrast distortion and they are represented as follows:

$$MSE = \frac{\sum_{x=1}^M \sum_{y=1}^N (\hat{f}(x, y) - f(x, y))^2}{M \times N} \tag{7}$$

$$RMSE = \sqrt{MSE} \tag{8}$$

$$PSNR = 10 \log_{10} \left(\frac{1}{MSE} \right) db \tag{9}$$

$$IQI = \frac{\sigma_{f\hat{f}}}{\sigma_f \sigma_{\hat{f}}} \cdot \frac{2\overline{f\hat{f}}}{(\overline{f})^2 + (\overline{\hat{f}})^2} \cdot \frac{2\sigma_f \sigma_{\hat{f}}}{\sigma_f^2 + \sigma_{\hat{f}}^2} \tag{10}$$

RESULTS

MATLAB R2009a was used to execute the existing order statistics filters like min, max, median and alpha trimmed mean filter. A mask size of 5*5 and 7*7 was used for all the filters. Fig. 1 and Fig.2 are the digital and OCT images which contain speckle noise. Fig. 3 to Fig. 10 represents the denoised digital images. Fig. 11 to Fig. 18 represents the denoised OCT images. The digital image and the OCT images that are used for the experiments are two dimensional images. Table 1 gives the comparison between the denoised digital image and the denoised OCT image through the image metrics root mean square error (RMSE), peak signal to noise ratio (PSNR) and image quality index IQI. Fig 19 shows the comparison of the order statistics filters with respect to the image metrics and it indicates that calculated RMSE values of OCT images are less when compared with digital images. However as the PSNR values for OCT images are high but the filtered OCT images are smoothed. As IQI values for the OCT images are less, it indicates that the order statistics filters reduces the contrast level of the filtered image.

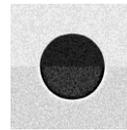


Fig. 1 Digital Image



Fig. 2 OCT Image

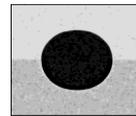


Fig. 3 min [5*5]

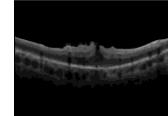


Fig. 11 min [5*5]

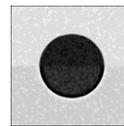


Fig. 4 max [5*5]



Fig. 12. max [5*5]

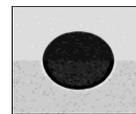


Fig. 5 median[5*5]



Fig. 13 median[5*5]



Fig. 6 ATM[5*5]



Fig. 14 ATM[5*5]

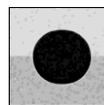


Fig. 7 min[7*7]

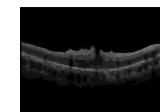


Fig. 15 min[7*7]



Fig. 8 max[7*7]

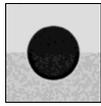


Fig. 9 median[7*7]

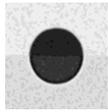


Fig. 10 ATM[7*7]

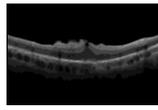


Fig. 16 max[7*7]

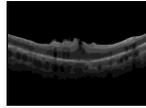


Fig. 17 median[7*7]



Fig. 18 ATM[7*7]

	Order Statistics Filters					
	Digital Image			OCT Image		
	RMSE	PSNR	IQI	RMSE	PSNR	IQI
Min [5*5]	13.435	25.57	0.8331	8.5517	29.49	0.3313
Min [7*7]	14.153	25.11	0.7839	9.0977	28.95	0.1999
Max [5*5]	8.196	29.86	0.9889	5.5468	33.25	0.4677
Max [7*7]	10.681	27.56	0.9407	7.1914	30.99	0.3478
Med [5*5]	10.362	27.82	0.952	6.8266	31.45	0.3927
Med [7*7]	12.032	26.52	0.8936	7.9139	30.16	0.3191
ATM [5*5]	15.885	24.11	0.9781	10.42	27.77	0.6394
ATM [7*7]	16.785	23.09	0.9726	11.54	25.65	0.1207

Table 1: Calculated values of image metrics like RMSE, PSNR and IQI for both digital and OCT image.

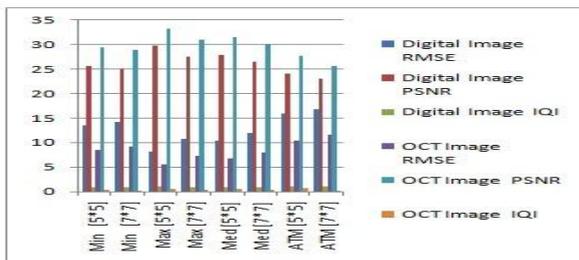


Fig. 19 Comparison between digital and OCT image.

V. CONCLUSION

In this work a digital image and a OCT image with speckle noise was used. The existing order statistics filters were used for removing the speckle noise. The results are evaluated through the image metrics like root mean square error, peak signal to noise ratio and image quality index. Through this work it is observed that the choice of filters for de-noising the medical images depends on the type of noise and type of filtering techniques.

REFERENCES

- [1] Brett E.Bouma, Guillermo J Tearney, Handbook of Optical coherence Tomography, NewYork, Marcell Dekker 2002.
- [2] A. Stella, Dr. Bhushan Trivedi, A Review on the use of Optical Coherence Tomography in Medical Imaging, International Journal of Advanced Research in Computer Science, volume 2, No 1, Jan-Feb 2011, 542-544.
- [3] H. Xie, L. E. Pierce, and F. T. Ulaby, Statistical properties of logarithmically transformed speckle, IEEE Transactions of Geoscience Remote Sensing, vol. 40, pp. 721-727, Mar. 2002.
- [4] R.C.Gonzalez and R.E. Woods, Digital Image Processing. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 2002.
- [5] A.K. Jain, Fundamentals of Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall; 1989.
- [6] H. G. Longotham and A.C. Bovik, Theory of order statistics filters and their relationship to linear FIR filters, IEEE Transactions on Acoustic. Speech, Signal Processing, vol. ASSP-37, no.2, pp. 275-287, February 1989.
- [7] S.Sridhar, Digital Image Processing, Oxford University Press, 2011.
- [8] P.G.J Barten, Contrast sensitivity of human eye and its effects on image quality, SPIE, Washington, 1999.
- [9] M. Marta, S. Grgic, M. Grgic, Picture quality measures in image compression systems, Proceedings EUROCON '03, p. 233-7, 2003.
- [10] Z. Wang, A.C. Bovik, "A universal image quality index", IEEE Signal Processing Letters, vol. 9, no. 3, pp.81-84, 2002.

Stochastic Model for Expected Time to Green House Effect

S. Vijaya

Faculty of Mathematics, Annamalai University, India.

ABSTRACT: One of the important aspects in the study of global warming is relating to increase of temperature. The factors like CO₂, CO and Nitrogen etc., plays a vital role to hasten the process of increase in global temperature. The only source of global warming is CO₂ emission. The interarrival times between two successive CO₂ emissions in this identified is a potential cause.[3] have obtained the expected time to Seroconversion and its variance when the interarrival times are identically independent random variables and also the case where they are correlated. In this paper the expected time to Green House effect is derived assuming the interarrival times are not independent.

Key words: Greenhouse effect, Global warming, threshold. The AMS classification number is 92C60.

I. INTRODUCTION

Stochastic models are widely used in the study of global warming and its consequences. There are many aspects which are taken for study the time to green house effect depending upon the increase of global temperature. If the global temperature crosses the threshold level which in turn leads to greenhouse effect. In developing such stochastic models the authors have used the concept of shock models and cumulative damage process discussed by [1]. [3] have obtained the expected time to Seroconversion under the assumption that the interarrival time between contacts are not independent but constantly correlated. In this paper it is assumed that the interarrival times which form a sample of observations that form order statistics and so they are not independent.

II. ASSUMPTION OF THE MODEL

1. By burning of fossil and other fuels, a random amount of CO₂ emission occurs.
2. CO₂ emission is the only source of global warming.
3. CO₂ emission is a damage process which is linear and cumulative
4. Increase of global temperature is caused by CO₂ emission are assumed to be identically independent random variable
5. If the global temperature exceeds threshold level Y which is itself is a random variable, then greenhouse effect takes place.
6. The process which generate the CO₂ emission, the sequence of increase in global temperature and threshold are mutually independent.
7. From the large number of CO₂ emissions between successive events, a random sample of K observations are taken

III. NOTATIONS

X_i - A random variable representing the increase of the global temperature due to CO₂ emission in the i^{th} event. X_i 's are i.i.d with p.d.f $g(\cdot)$ and distribution function $G(\cdot)$.
 Y - A random variable representing the global warming threshold with p.d.f $h(\cdot)$ and distribution function $H(\cdot)$.
 U_i - A random variable representing the interarrival times between successive events, $i = 1, 2, \dots, k$, with p.d.f $f(\cdot)$.
 $g_k(\cdot)$ - The p.d.f. of the random variable $\sum X_i$, $i = 1, 2, \dots, k$.
 $F_k(\cdot)$ - The k th convolution of $F(\cdot)$.
 T - The continuous random variable denoting the time to Green house effect.
 $U_{(i)}$ - Smallest order statistic with p.d.f $f_{u(1)}(t)$.
 $U_{(k)}$ - The largest order statistic with p.d.f $f_{u(k)}(t)$.
 $f^*(s)$ - Laplace transform of $f(\cdot)$.
 $f_{u(1)}^*(s), f_{u(k)}^*(s)$ - Laplace transform of $f_{u(1)}(\cdot)$ and $f_{u(k)}(\cdot)$ respectively.

IV. RESULTS

If X_i $i = 1, 2, \dots, k$ are the contributions to the global warming in k events during the period $(0, t)$, the time to cross the global warming threshold level can be obtained as follows.

$S(t)$ - Survivor function = $P(T > t)$

$$P(T > t) = \sum_{K=0}^{\infty} \Pr[\text{there are exactly } k \text{ events on } (0, t)] \\ * \Pr[\text{the cumulative total of global warming} < Y]$$

It can be shown that

$$P\left[\sum_{i=1}^k X_i < Y\right] = \int_0^{\infty} g_k(x) H(x) dx$$

Assuming that $y \sim \exp(\theta)$, we have

$$P[\sum_{i=1}^k X_i < Y] = \int_0^{\infty} g_k(x) e^{-\theta x} dx = g_k^*(\theta) = [g^*(\theta)]^k \quad \text{-----(4.1)}$$

since x_i are all i.i.d,

Also P_r [exactly k shocks in $(0,t)$] = $F_k(t) - F_{k+1}(t)$
 [by renewal theory]

$$\begin{aligned} \text{Hence } P(T > t) &= \sum_{k=0}^{\infty} [F_k(t) - F_{k+1}(t)] [g^*(\theta)]^k \\ &= 1 - [1 - g^*(\theta)] \sum_{k=1}^{\infty} F_k(t) (g^*(\theta))^{k-1} \text{ on simplification} \end{aligned}$$

$$\begin{aligned} \text{Hence } L(t) &= P[T < t] = 1 - S(t) \\ &= [1 - g^*(\theta)] \sum_{k=1}^{\infty} F_k(t) [g^*(\theta)]^{k-1} \quad \text{----- (4.2)} \end{aligned}$$

Now taking the Laplace transform of $L(t)$, it can shown that

$$L^*(s) = \frac{[1 - g^*(\theta)] f^*(s)}{1 - g^*(\theta) f^*(s)} \quad \text{-----(4.3)}$$

on simplification

The interarrival times U_1, U_2, \dots, U_k are i.i.d random variables and $U_{(1)} < U_{(2)} < \dots < U_{(k)}$ form k order statistics which are also random variables which are not independent.

Now the p.d.f of $U_{(k)}$ is

$$F_{U_{(k)}}(t) = k[F(t)]^{k-1} f(t) \quad \text{-----(4.5)}$$

Assuming that $f(t) \sim \exp(\lambda)$, it can be shown that

$$F_{u_{(k)}}^*(s) = \frac{k! \lambda^k}{(\lambda+S)(2\lambda+S) \dots (k\lambda+S)} \quad \text{----- (4.6)}$$

Substituting (4.6) in (4.3) and assuming $g(\cdot) \sim \exp(\theta)$ it can be shown that

$$\square^*(s) = \frac{\theta k! \lambda^k}{(c+\theta)(\lambda+s)(2\lambda+s) \dots (k\lambda+s) - ck! \lambda^k} \quad \text{-----(4.7)}$$

on simplification.

Now the expected time to greenhouse effect is given by $E(T) = \frac{-d \square^*(s)}{ds}$ at $s=0$.

$$= \frac{c + \theta}{c\lambda} \sum_{n=1}^k \frac{1/n}{\lambda^2 \theta} \quad \text{-----(4.8)}$$

It implies that $E(T)$ becomes larger as k increases since $\sum 1/n$ increases with K . Hence it may be concluded that the maximum value $U_{(k)}$ increases with increase in k and also the interarrival time $U_{(k)}$ becomes longer. This results in the larger value of $E(T)$.

It may be shown that

$$\begin{aligned} E(T^2) &= \frac{d^2 \square^*(s)}{ds^2} \Big|_{s=0} \\ &= \frac{2(c+\theta)^2}{\theta^2 \lambda^2} \left[\sum_{n=1}^k \frac{1/n}{\lambda^2 \theta} \right]^2 + \frac{c+\theta}{\lambda^2 \theta} \left[\sum_{n=1}^k \frac{1/n^2}{\lambda^2 \theta} \right] \quad \text{-----(4.9)} \end{aligned}$$

We have

$$\begin{aligned} V(T) &= \text{Variance of } T \\ &= E(T^2) - [E(T)]^2 \\ &= \frac{c(c+\theta)}{\theta^2 \lambda^2} \sum_{n=1}^k \frac{1/n}{\lambda^2 \theta} + \frac{c+\theta}{\lambda^2 \theta} \sum_{n=1}^k \frac{1/n^2}{\lambda^2 \theta} > 0 \quad \text{-----(4.10)} \end{aligned}$$

This implies that for fixed C, λ, θ the variance of T increases as k increases.

Now considering the first order statistics $U_{(1)}$ it can be shown that

$$f_{U_{(1)}}^*(s) = \frac{k\lambda}{\lambda k + s} \quad \text{-----(4.11)}$$

Substituting (4.11) in (4.3) for $f^*(s)$ it is seen that

$$\phi^*(s) = \frac{c \lambda k}{((c+\theta)(\lambda k + s) - (\theta \lambda k))} \quad \text{-----(4.12)}$$

So that the mean and variance of T are obtained as

$$E(T) = \frac{-d \phi^*(s)}{ds} \Big|_{s=0} = \frac{C+\theta}{C\lambda k} \quad \text{-----(4.13)}$$

$$V(T) = \text{Variance of T} = E(T^2) - [E(T)]^2 = \frac{2(C+\theta)^2}{C^2\lambda^2k^2} - \frac{(C+\theta)^2}{C^2\lambda^2k^2} = \left(\frac{C+\theta}{C\lambda k} \right)^2 \quad \text{-----(4.14)}$$

V. NUMERICAL ILLUSTRATION

For the case of $U_{(k)}(t)$

C = 1.5 θ = 0.5 λ = 1.0 are fixed

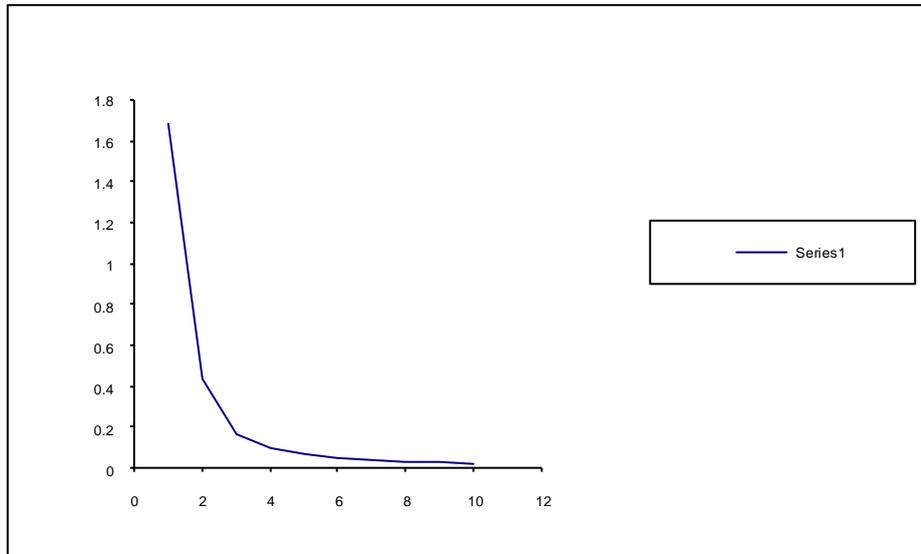
k	E(T)	V(T)
1	4.0	16
2	6.0	32
3	7.32	45.4
4	8.32	57.59
5	9.12	68.21
6	9.76	77.348
7	10.32	85.92
8	10.8	93.56
9	11.24	100.87
10	11.64	107.77

For the case of $U_{(1)}(t)$

$C=1.5$	$\theta=0.5$	$\lambda=1.0$ are fixed
k	$E(T)$	$V(T)$
1	1.3	1.69
2	0.6	0.435
3	0.4	0.16
4	0.3	0.09
5	0.26	0.067
6	0.22	0.048
7	0.19	0.036

$C = 1.5$ $\theta = 0.5$ $\lambda = 1.0$

For the case of $U_{(1)}(t)$ Variance



VI. CONCLUSION

It is very interesting to observe the following from the study of the numerical rate and the respective graphs.

1. The values of $E(T)$ and $V(T)$ both increase with an increase in 'k', namely the number of events. If k becomes larger than the corresponding $U_{(k)}$ also becomes larger thereby implying that it is the largest of the interarrival times. In such a case the inter global warming times are elongated thereby having a delayed time to greenhouse effect. Hence the curves for $E(T)$ and $V(T)$ go upwards in both the cases.
2. The values of $E(T)$ and $V(T)$ both decreases with an increase in 'k' for the case of $U_{(1)}(t)$. If the inter global warming time is the smallest random variable, than it implies that if number of events are more, the greenhouse effect will be much earlier and the process gets speeded up resulting in a decline in both of $E(T)$ and $V(T)$ as indicated by the graphs.

REFERENCES

- [1] Esary, J.D., Marshall, A.W. and Proschan, F. (1973). Shock models and wear processes. *Ann. Probability*, 627-649.
- [2] Stilianakis, N., Schenzle, D. and Dietz, K.(1994). On the antigenic diversity threshold model for AIDs. *Mathematical Biosciences*, 121, 235-247.
- [3] Sathiyamorthi, R. and Kannan, R. (1998). On the time to Sero-conversion of HIV patients under correlated inter contact times. *Pure and Applied Mathematika Sciences*, Vol. XLVIII, No. 1-2.

Survey of Different Biometrics Techniques

Jitendra Choudhary

Asst.Prof.Computer Science &Engineering Department, IES College Bhopal, India

ABSTRACT: *Biometrics is the study of automatically recognizing humans by means of inherently unique physical or behavioural characteristics. The human physical characteristics like fingerprints, face, hand geometry, voice, Signature, palm and iris are known as biometrics. These features are used to provide an authentication for computer based security systems. The existing computer security systems used at various places like banking, passport, credit cards, smart cards, PIN, access control and network security are using username and passwords for person identification. Biometric systems also introduce an aspect of user convenience. For example, they alleviate the need for a user to remember multiple passwords associated with different applications. A biometric system that uses a single biometric trait for recognition has to contend with problems related to non-universality of the trait, spoof attacks, limited degrees of freedom, large intra-class variability, and noisy data. Some of these problems can be addressed by integrating the evidence presented by multiple biometric traits of a user (e.g., face and iris). Such systems, known as multimodal biometric systems, In this paper, the main focus is on the various biometrics Techniques, their Types applications and the biometrics recognition systems.*

Keywords: *Biometrics, computer based security systems, Feature Extraction, Biometrics recognition system, false reject rate, false accept rate.*

I. INTRODUCTION

Biometrics is the science and technology of measuring and analyzing biological data. In information technology, biometrics refers to technologies that measure and analyze human body characteristics, such as DNA, fingerprints, eye retinas and irises, voice patterns, facial patterns and hand measurements, for authentication purposes.

1.1 Physical biometrics

These involve some form of physical measurement and include modalities such as face, fingerprints, iris-scans, hand geometry etc.

1.2. Behavioural biometrics

These are usually temporal in nature and involve measuring the way in which a user performs certain tasks. This includes modalities such as speech, signature, gait, keystroke dynamics etc.

1.3. Chemical biometrics

This is still a nascent field and involves measuring chemical cues such as order and the chemical composition of human perspiration A number of biometric identifiers are in use in various applications. Each biometric has its strengths and weaknesses and the choice typically depends on the application.

II. TAXONOMY OF BIOMETRICS

According to biometrics applications and technologies, we can mainly classify biometrics systems into the following two types-

2.1. Application Type

As the general knowledge, biometrics technology is basically applied to security services infect, a few other applications using biometrics are also very effective based on a summary of different applications, biometrics system can be divided into four categories.

2.2. Personal Authentication

We may use biometrics technology to identify individual since this is popular biometrics application.

2.3. Medical Diagnosis

Tongue, color of face beat of heart and other aspects of our body can be also used as biometrics features for medical diagnosis. Traditional chine's medicine (TCM) particularly needs to measure such kind of body characteristics.

2.4. Future Expectation

Egypt and china have some specialists on this biometrics application. by looking at one's palm, the specialists can tell the personality as well as the future direction of a person.

2.5. Technology exploration

Measuring body characteristics can also used to decide one's ethic .we may use this biometrics technology to monitor the population shifting among different areas.

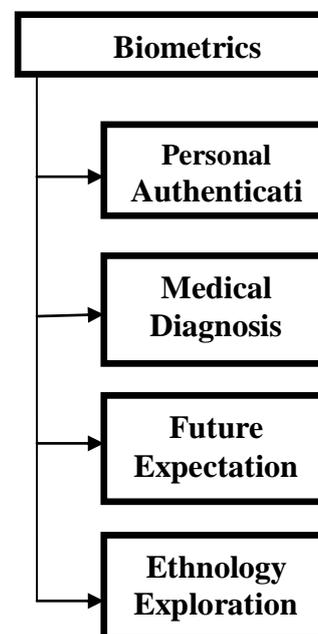


Fig. 1: Taxonomy by Application Type

III. TECHNOLOGY TYPE

System are categorized in terms of the employed physical or behavioral characteristics type, biometrics system are categorized in terms of the employed physical or behavioral characteristics. Since we will emphasize on the application of personal authentication.

No single biometric is expected to effectively meet the requirements of all the applications. The match between a biometric and an application is determined depending upon the characteristics of the application and the properties of the biometric. These various biometric identifiers mentioned above are compared in table below fingerprint recognition has a very good balance of all the desirable properties.

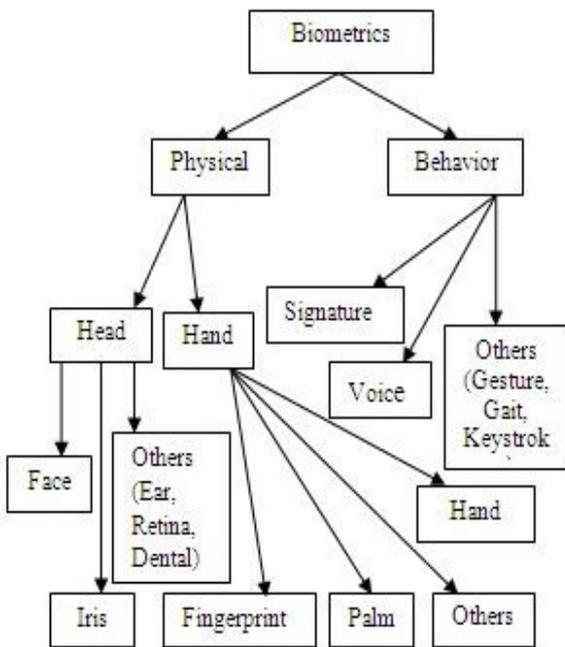


Fig.2: Taxonomy by Technology Type.

3.1. Description of the different types of biometrics

Biometrics can be physiological or behavioural. Physiological biometrics is more common and accurate than behavioural. Some of physiological biometrics, that was the most commonly used in many fields, is facial scan, iris scan, hand scan and fingerprint scan. However keystroke, signature and voice recognition are some of behavioural biometrics. A brief introduction of some of this biometrics, are described below [1].

3.2. Face

Facial features are the most normal feature used by human to recognize one another. Face recognition is based on both the shape and location of the eyes, eyebrows, nose, lips, and chin or on the overall analysis of the face image that represent a face as a number of recognized faces. In a face recognition system, it is hard to match face images taken from two different views and under different illumination conditions. Moreover the face of individual can be changed by times. All this criteria make face recognition system to be uncertain if really the face itself is enough to recognize a person from a large number of identities [3].



Fig.3: Facial image variations amongst the same subject

3.3. Hand geometry

This method is simple, easy and inexpensive. It has been established in many locations in the world. Hand geometry recognition systems are based on a number of measurements taken from the human hand. It measures the shape of hand, the size of palm, and the lengths and widths of the fingers. Many environmental or anomalies factors don't affect any change in the accuracy of this recognition system. However this system can't be generalized to recognize from a large number of population. In addition the geometry of the hand is not constant; it can be changed in proportion of the age. Moreover the size of the hand is big and it is not currently in wide deployment for computer security applications primarily because it requires a large scanner [2].



Fig.4: Commercial three-dimensional scanner

3.4. Iris

Iris recognition is based on the features that exist in the coloured tissue surrounding the pupil which has many interesting points that can be used for comparison, including rings, rows and spots [4]. The texture of the iris is very complex and distinctive which is very useful for the recognition system. Even the irises of identical twins are different. Although based on this complexity and this distinctness, the system is more accurately deployed and supports the probability of extensive identification systems [12].

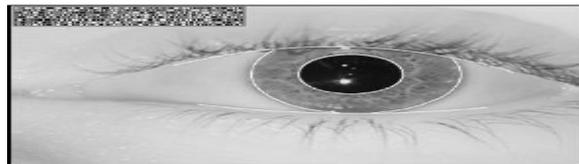


Fig.5: Example of an iris pattern.

3.5. Keystroke

The way and the manner of typing on computer keyboard vary from individual to individual. This biometric is not considered as unique but it can be sufficient for some applications. Identification of this behavioural biometrics is basically what a person types is less important than how he types it. Using this approach several things can be analyzed: time between key-pressed and key-released, type of keyboard used or the emotional and physical state of the

person. So no special hardware is required for keystroke analysis, just the usual computer keyboard.

3.6. Signature

Signature is the way a person signs his name. Depending on this sign, the individual can be identified. Signature recognition examines the unique way in which the signature is written. In the signature recognition system, the signature is compared by examining how the signature was written or it is verified by estimating how the signature was created. Sometimes, this type of biometric can be changed over time. The same person can sign in different way. In addition, it is affected by physical conditions such as sickness or sentimental condition such as individual's feeling.



Fig .6: Signature biometrics

3.7. Voice

Voice recognition is the identification of a person based on unique characteristic on their voice. Voice characteristic is the combination between physical and behavioural biometric. For the physical part of view, voice is constant because it depends on the size or shape of the mouth, lips, vocal tracts and nasal cavities and so on. However for the behavioural part, voice is not constant. It can be changed based on individual's emotion, sickness or age. Due to this behavioural effect, voice recognition system can't not be considered as a distinctive biometric [6].



Fig .7: Voiceprint

3.8. Gait

Gait recognition is a particular type of biometric due to its capability to identify a person at distance. Gait is related to the way of the person walking. The gait recognition system use standard camera in any conditions and develop algorithms to extract the silhouette of the person in case he is moving. Therefore the system can track the person over time. However the algorithm is not very efficient for this trait is affected by many conditions such as the type of cloth's or shoes the individual's wearing, the walking surface or the health. All these biometrics are acceptable in different environment and none of them is optimal. However the most accurate ones are iris and fingerprint techniques. Due to the fact iris recognition is expensive and it requires advance requirement, fingerprint is one of the most mature biometrics and suitable for many applications [5]. Fingerprint biometrics is very distinctive, not expensive, unique and permanent and has a very good balance from all the properties.

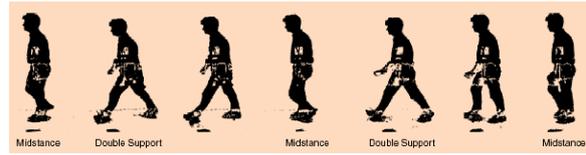


Fig .8: Samples recorded from a gait cycle.

3.9. Fingerprint

Fingerprint recognition system is the oldest recognition system among all the biometrics techniques. Everyone have a unique and unchangeable fingerprint. Like iris, fingerprints of the twins are even different. Based on this uniqueness and distinctness, fingerprint identification is used in many applications for a long period. A fingerprint is the pattern of ridges and valleys on the surface of the finger. It can be changeable only by some environmental and job-related factors such as cuts or injuries on the finger. These factors cause to be the system unsuitable in some degree. Generally the accuracy of the fingerprint recognition is sufficient in many applications especially in Forensics. To allow great identification systems for a large number of identities, the systems require having a multiple fingerprint from the same person to give additional information [7].



Fig.9: Fingerprint sensors in everyday products.

IV. ENROLLMENT AND AUTHENTICATION

In enrollment the biometrics of the user is captured and the extracted features templates are stored in the database. In authentication the biometrics of the user is captured again and the extracted features are compared with the ones already existing in the database to determine a match. The specific record to fetch from the database is determined using the claimed identity of the user. The database itself may be central or distributed with each user carrying his template on a smart card.

Biometrics is the science of verifying the identity of an individual through physiological measurements or behavioural traits. Since biometric identifiers are associated permanently with the user they are more reliable than token or knowledge based authentication methods. Biometrics offers several advantages over traditional security measures [8]. These includes

1. Non-repudiation
2. Accuracy and Security
3. Screening

4.1. Non-repudiation

With token and password based approaches, the perpetrator can always deny committing the crime pleading that his/her password or ID was stolen or compromised even when confronted with an electronic audit trail. There is no way in which his claim can be verified effectively. This is known as the problem of deniability or of 'repudiation'. However, biometrics is indefinitely associated with a user and hence it cannot be lent or stolen making such repudiation infeasible.

4.2. Accuracy and Security

Password based systems are prone to dictionary and brute force attacks. Furthermore, such systems are as vulnerable as their weakest password. On the other hand, biometric authentication requires the physical presence of the user and therefore cannot be circumvented through a dictionary or brute force style attack. Biometrics has also been shown to possess a higher bit strength compared to password based systems and is therefore inherently secure.

4.3. Screening

In screening applications, we are interested in preventing the users from assuming multiple identities e.g. a terrorist using multiple passports to enter a foreign country. This requires that we ensure a person has not already enrolled under another assumed identity before adding his new record into the database. Such screening is not possible using traditional authentication mechanisms and biometrics provides the only available solution. The various biometric modalities can be broadly categorized as the various stages of a typical fingerprint recognition system are shown in Figure 4.3 the fingerprint image is acquired using off-line methods such as creating an inked impression on paper or through a live capture device consisting of an optical, capacitive, ultrasound or thermal sensor. The first stage consists of standard image processing algorithms such as noise removal and smoothening.

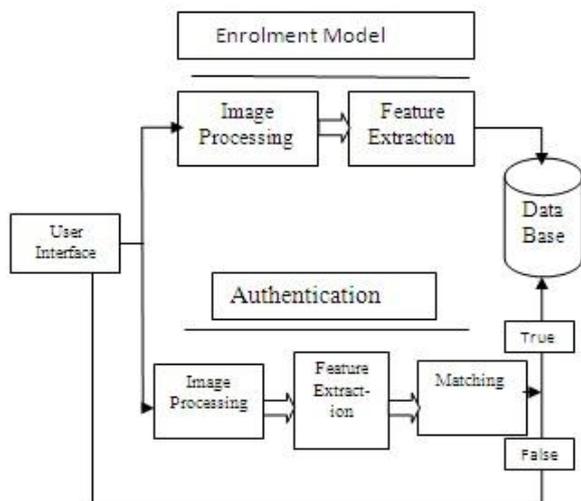


Fig.10: General processes to recognize a fingerprint

However, it is to be noted that unlike regular images, the fingerprint image represents a system of oriented texture and has very rich structural information within the image. Furthermore, the definition of noise and unwanted artifacts are also specific to fingerprints. The fingerprint image enhancement algorithms are specifically designed to exploit the periodic and directional nature of the ridges. Finally, the minutiae features are extracted from the image and are subsequently used for matching. Although research in fingerprint verification research has been pursued for several decades.

4.4. Biometrics and Pattern Recognition

As recently as a decade ago, biometrics did not exist as a separate field. It has evolved through interaction and confluence of several fields. Fingerprint recognition

emerged from the application of pattern recognition to forensics. Speaker verification evolved out of the signal processing community. Face detection and recognition was largely researched by the computer vision community. While biometrics is primarily considered as application of pattern recognition techniques, it has several outstanding differences from conventional classification problems as enumerated below

1. In a conventional pattern classification problem such as optical character recognition (OCR) recognition the number of patterns to classify is small (A-Z) compared to the number of samples available for each class. However in case of biometric recognition, the number of classes is as large as the set of individuals in the database. Moreover it is very common that only a single template is registered per user.

2. The primary task in biometric recognition is that of choosing a proper feature representation. Once the features are carefully chosen, the act of performing verification is fairly straight-forward and commonly employs simple metrics such as Euclidean distance. Hence the most challenging aspects of biometric identification involve signal and image processing for feature extraction.

3. Since biometric templates represent personally identifiable information of individuals, security and privacy of the data is of particular importance unlike other applications of pattern recognition.

4. Modalities such as fingerprints, where the template is expressed as an unordered point set minutiae do not fall under the category of traditional multi-variety features commonly used in pattern recognition.

4.5. Fingerprints as a Biometric

Fingerprints were accepted formally as valid personal identifier in the early twentieth century and have since then become a defect to authentication technique in law-enforcement agencies worldwide. The federal bureau of investigation (FBI) currently maintains more than 400 million fingerprint records on file. Fingerprints have several advantages over other biometrics such as the following.

1. High universality
2. High distinctiveness
3. High permanence
4. Easy collectability
5. High performance
6. Wide acceptability

V. BIOMETRIC RECOGNITION SYSTEM

The Biometric Recognition Systems are used to identify the person based on the feature vectors of any one of the biometric that the person possesses [10]. These systems are person authorized systems hence offer more secure and convenient process of identification compared to alternative methods of identification. The biometric

System uses the individual's physical characteristics like fingerprint, hand geometry, face, voice or iris. They are more reliable and secure as they provides the access to authorized users in their physical presence [11].

5.1. Fingerprint Recognition Systems

Fingerprint recognition systems consist of the following parts

- Sensing or Image acquisition
- Pre-processing
- Feature or minutiae extraction
- Matching

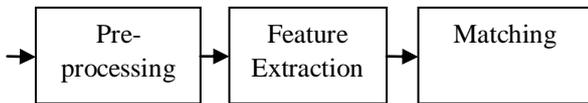


Fig.11: Fingerprint Recognition Systems

5.2. Fingerprint matching techniques

The large number of approaches to fingerprint matching can be coarsely classified into three families.

5.2.1. Correlation-based matching

Two fingerprint images are superimposed and the correlation between corresponding pixels is computed for different alignments (e.g. various displacements and rotations).

5.2.2. Minutiae-based matching

This is the most popular and widely used technique, being the basis of the fingerprint comparison made by fingerprint examiners. Minutiae are extracted from the two fingerprints and stored as sets of points in the two-dimensional plane. Minutiae-based matching essentially consists of finding the alignment between the template and the input minutiae sets that result in the maximum number of minutiae pairings

5.2.3. Pattern-based (or image-based) matching

Pattern based algorithms compare the basic fingerprint patterns (arch, whorl, and loop) between a previously stored template and a candidate fingerprint. This requires that the images be aligned in the same orientation. To do this, the algorithm finds a central point in the fingerprint image and centers on that. In a pattern-based algorithm, the template contains the type, size, and orientation of patterns within the aligned fingerprint image. The candidate fingerprint image is graphically compared with the template to determine the degree to which they match.

5.3. The metrics of Biometrics

- FTE – Failure To Enroll
- FTA – Failure To Accept
- FAR – False Acceptance Rates
- FRR – False Reject Rates

5.4. Essential parameters

- Livens testing
- Tamper resistance
- Secure communication
- Security Threshold level
- Fall back node

5.5. Fingerprint techniques

- Optical
- Capacitive
- Thermal
- Ultrasonic



Fig.12: Fingerprint Image

VI. BIOMETRIC CHARACTERISTIC

Every biometric characteristic require satisfying the following properties:

Universality: The characteristic should be encountered in each person.

Uniqueness: The characteristic should be unique between individuals.

Permanence: The characteristic should be sufficiently constant over long time.

Collect ability: The characteristic can be measured from a quantity point of view.

Performance: The characteristic give an accurate result under different environment.

Acceptability: The entire people accept to give their traits to the system without any Problem.

Circumvention: The characteristic must be hard to deceive and imitate.

6.1. Feature Extraction Module

In the feature extraction module, the preprocessed image voice is used to extract the features. The feature extraction algorithms are applied to get feature vector of the biometric image / voice. There are various feature extraction techniques like Independent Component Analysis, Linear discriminate component, principal component analysis, wavelet transform, LPC, MFCC, etc [13][14][15][16]. According to the biometrics selected and its application the feature extraction technique can be applied.

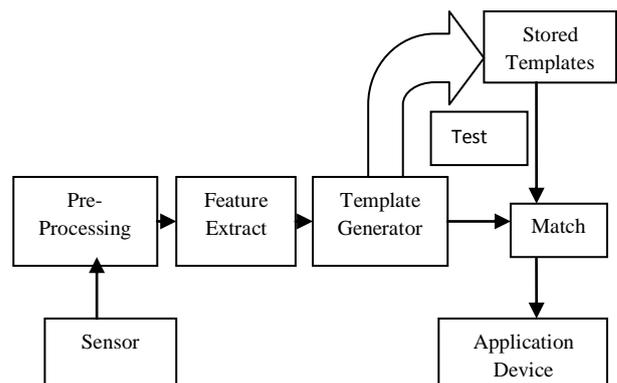


Fig.13: Biometric System

VII. APPLICATIONS OF THE BIOMETRIC RECOGNITION SYSTEMS

Which uses the physical biological or behavioral characteristics that can be processed to perform automatic recognition of a person Hence this requires achieving low cost, reliable human identification system by using feature set of individual characteristics. The biometric concentrates on physical aspects like finger print, hand geometry, face, voice and iris of a person [12].

In the last years has considerably increased the area of application of biometrics and it's expected that in the near future, we will use biometry many times in our daily activities such as getting in the car, opening the door of our house, accessing to our bank account, shopping by internet, accessing to our PDA, Mobil phone, laptops, etc. Depending of where the biometrics is deployed, the applications can be categorized in the following five main groups: forensic, government, commercial, health-care and traveling and immigration. However, some applications are common to these groups such as physical access, PC/network access, time and attendance, etc.

VIII. CONCLUSION

Biometric recognition, or biometrics, refers to the automatic identification of a person based on his/her anatomical (e.g., fingerprint, iris) or behavioral (e.g., signature) characteristics or traits. This method of identification offers several advantages over traditional methods involving ID cards (tokens) or PIN numbers (passwords) for various reasons: (i) the person to be identified is required to be physically present at the point-of-identification; (ii) identification based on biometric techniques obviates the need to remember a password or carry a token. With the increased integration of computers and Internet into our everyday lives, it is necessary to protect sensitive and personal data. By replacing PINs (or using biometrics in addition to PINs), biometric techniques can potentially prevent unauthorized access to ATMs, cellular phones, laptops, and computer networks. The biometric features can be easily.

Acquired and measured for the processing only in the presence of a person. Hence these systems are proved highly confidential computer based security systems.

REFERENCES

- [1] Faundez-Zanuy, M. Biometric verification by means of hand geometry. 39th Annual 2005 International Carnahan Conference on Security Technology, 61-67(2005).
- [2] F.A. Afsar, M. Arif and M. Hussain, "Fingerprint Identification and Verification System using Minutiae Matching", National Conference on Emerging Technologies, 2004.
- [3] Rafael C. Gonzalez, Richard E Woods and Steven L Eddins "Digital Image Processing using MATLAB", Pearson Education, 2004.
- [4] Sulochana Sonkamble, Dr. R.C. Thool, Balwant Sonkamble, "An Effective Machine-Vision System for Information Security and Privacy using Iris Biometrics", in The 12th World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2008 at Orlando, Florida, USA during June 29th - July 2nd, 2008.
- [5] Boulgouris, N.V., Hatzinakos, D., &Plataniotis, K.N. Gait recognition: a challenging signal processing technology for biometric identification. IEEE Signal Processing Magazine, 22, 78-90(2005).
- [6] L. Rabiner, B. H. Juang, "Fundamentals of Speech Recognition", Pearson Education.
- [7] K. Karu, A.K. Jain, "Fingerprint classification, Pattern Recognition", 1996.
- [8] Ton van der Putte and JeroenKeuning "Biometrical Fingerprint Recognition", Fourth Working Conference on Smart Card Research and Advanced Applications, pp 289-303, 2000.
- [9] F.A. Afsar, M. Arif and M. Hussain, "Fingerprint Identification and Verification System using Minutiae Matching", National Conference on Emerging Technologies, 2004.
- [10] K.V.Kale, R.R.Manza, V.T.Humbe, Prapti Deshmukh" fingerprint image enhancement using composite technique" f-14-14/2003(SR).
- [11] Jain, A., Hong, L., Pankanti, S., and Bolle, R. An identity authentication system using fingerprints. In Proceedings of the IEEE (September 1997), vol. 85, pp. 1365-1388.
- [12] John Carter, Mark Nixon, "An Integrated Biometric Database" Department of Electronics and Computer Science, University of Southampton, Highfield, Southampton, SO95NH.
- [13] Peng Wang; Qiang Ji; Wayman, J.L., "Modeling and Predicting Face Recognition System Performance Based on Analysis of Similarity Scores", Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 29, Issue 4, April 2007.
- [14] Steve Lawrence C. Lee Giles Ah Chung Tsoi, Andrew D. Back, "Face Recognition: A Convolutional Neural Network Approach", IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition.
- [15] Wei Han, Cheong- Fat Chan, Chiu Sing Choy and Kong Pang Pun, "An Efficient MFCC Extraction Method in Speech Recognition", IEEE 2006.
- [16] John Daugman, "How Iris Recognition Works", IEEE transactions on circuits and systems for video technology, vol. 14, no. 1, January 2004.

An Oscillation Damping Reduced in Transmission line by Using Improve the Series-Shunt Controller

Azmeera Rajesh¹, S. Anupama²

^{1,2}(Department of Electrical and Electronic Engineering, Annamacharya Institute of Technology And Sciences
Andrapradesh, India)

Abstract: This paper proposes a new control approach for a unified power flow controller (UPFC) for power system oscillation damping. This control is simple to implement, yet is valid over a wide range of operating conditions. It is also effective in the presence of multiple modes of oscillation. The proposed control is implemented in several test systems and is compared against a traditional PI control.

Index Terms: Oscillation damping, power system stability, unified power flow controller (UPFC).

I. INTRODUCTION

ONE of bulk the most promising network controllers for the power system are the family of power electronics based controllers, known as “flexible ac transmission system” (FACTS) devices. FACTS devices work by modifying power flow in individual lines of the power grid, maintaining voltage stability, and damping oscillations. The DOE National Transmission Grid Study released in May 2002 identified FACTS devices as playing a significant role in the “Intelligent Energy System” of the future. This rapid control has been shown to be effective in achieving voltage support and stability improvement, thus allowing the transmission system to be operated more efficiently with a smaller stability margin. The rapid development of the power electronics industry has made FACTS devices increasingly attractive for utility deployment due to their flexibility and ability to effectively control power system dynamics. The primary function of the FACTS is to control the transmission line power flow; the secondary functions of the FACTS can be voltage control, transient stability improvement and oscillation damping. The unified power flow controller (UPFC) is the most versatile FACTS device. The UPFC is able to simultaneously provide both series and shunt compensation to a transmission line providing separate control of the active and reactive powers on the transmission line.

In recent years, the use of the UPFC for oscillation damping has received increased attention. Several approaches have been taken to the modeling and control of the UPFC. Perhaps the most common approach is to model the UPFC as a power injection model. The power injection model neglects the dynamics of the UPFC and uses the UPFC active and reactive power injection as the control inputs into the power system. This approach has the advantages of simplicity and computational efficiency since the fast dynamics of the UPFC are neglected.

While effective from a high-level control viewpoint, this approach assumes that the UPFC is ideally (and instantaneously) able to provide the required active and reactive powers.

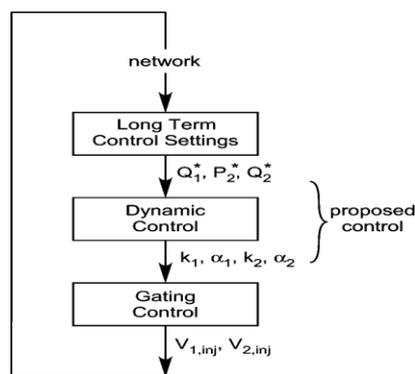


Fig. 1. UPFC hierarchical control.

In the cases where UPFC dynamics are included, the most common approach to controlling the UPFC has been to use PI control. PI control is simple to implement, yet very effective in damping an oscillatory mode when it is properly tuned. PI control is less effective in damping oscillations that contain multiple modes. For multiple mode damping, several lead-lag blocks are required that require additional coordinated tuning. Secondly, PI control becomes increasingly less effective as the system conditions move from the operating point around which the controller was tuned. For this reason, it is desirable to develop controllers that are impervious to changes in operating condition and that are effective for multiple mode

oscillation damping. In this paper, a new control is proposed that is effective over a wider range of operating conditions and is able to damp multiple modes effectively. Specifically in this paper, we will

- 1) Present the development of a new control for the UPFC;
- 2) Show its effectiveness over a range of operating conditions and multiple modes; and
- 3) Compare it with a traditional PI control method.

The proposed control is the dynamic control required to achieve given active and reactive power flow and voltage set-points. All system level control tacitly assumes that the UPFC can instantaneously achieve the required set points. However, there can be degradation of performance if a significant time lag exists between the time the set point is given and the time at which the UPFC achieves the desired injected voltages. The proposed controller is a nonlinear dynamic control that translates the desired system level control into gating control as shown in Fig. 1. Ideally, the system level control set points would not be constant power flows (or current injections), but rather would be

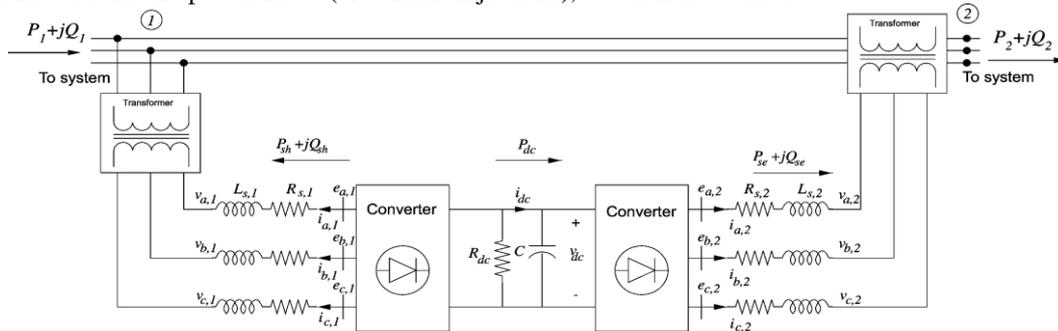


Fig. 2. Unified power flow controller diagram

Time varying in order to achieve the desired system response this paper asserts that a linear control is inadequate to track a moving target due to the requirement that it be tuned for different operating conditions and requires a much larger number of parameters. The proposed nonlinear control has the advantages of rapid tracking and is independent of tuning.

II. UPFC STATE MODEL

The unified power flow controller, or UPFC, is the most versatile FACTS device. It consists of a combination of a shunt and series branches connected through the dc capacitor as shown in Fig. 2. The series connected inverter injects a voltage with controllable magnitude and phase angle in series with the transmission line, therefore providing real and reactive power to the transmission line. The shunt-connected inverter provides the real power drawn by the series branch and the losses and can independently provide reactive compensation to the system. The UPFC model is a combination of the synchronous static compensator (STATCOM) and static series synchronous compensator (SSSC) models as follows

$$\frac{1}{\omega_s} \frac{d}{dt} i_{d1} = \frac{k_1 V_{dc}}{L_{s1}} \cos(\alpha_1 + \theta_1) + \frac{\omega}{\omega_s} i_{q1} - \frac{R_{s1}}{L_{s1}} i_{d1} - \frac{V_1}{L_{s1}} \cos \theta_1 \quad (1)$$

$$\frac{1}{\omega_s} \frac{d}{dt} i_{q1} = \frac{k_1 V_{dc}}{L_{s1}} \sin(\alpha_1 + \theta_1) - \frac{R_{s1}}{L_{s1}} i_{q1} - \frac{\omega}{\omega_s} i_{d1} - \frac{V_1}{L_{s1}} \sin \theta_1 \quad (2)$$

$$\frac{1}{\omega_s} \frac{d}{dt} i_{d2} = -\frac{R_{s2}}{L_{s2}} i_{d2} + \frac{\omega}{\omega_s} i_{q2} + \frac{k_2}{L_{s2}} \cos(\alpha_2 + \theta_1) V_{dc} - \frac{1}{L_{s2}} (V_2 \cos \theta_2 - V_1 \cos \theta_1) \quad (3)$$

$$\frac{1}{\omega_s} \frac{d}{dt} i_{q2} = -\frac{R_{s2}}{L_{s2}} i_{q2} - \frac{\omega}{\omega_s} i_{d2} + \frac{k_2}{L_{s2}} \sin(\alpha_2 + \theta_1) V_{dc} - \frac{1}{L_{s2}} (V_2 \sin \theta_2 - V_1 \sin \theta_1) \quad (4)$$

$$\begin{aligned} \frac{C}{\omega_s} \frac{d}{dt} V_{dc} = & -k_1 \cos(\alpha_1 + \theta_1) i_{d1} \\ & - k_1 \sin(\alpha_1 + \theta_1) i_{q1} - k_2 \cos(\alpha_2 + \theta_1) i_{d2} \\ & - k_2 \sin(\alpha_2 + \theta_1) i_{q2} - \frac{V_{dc}}{R_{dc}} \end{aligned} \quad (5)$$

where the parameters are as shown in Fig. 2. The currents i_{d1} and i_{q1} are the dq components of the shunt current. The currents i_{d2} and i_{q2} are the dq components of the series current. The voltages $V_1 \angle \theta_1$ and $V_2 \angle \theta_2$ are the shunt and series voltage magnitudes and angles, respectively. The UPFC is controlled by varying the phase angles (α_1, α_2) and magnitudes (k_1, k_2) of the converter shunt and series output voltages (e_1, e_2), respectively.

The power balance equations at bus 1 are given by

$$\begin{aligned} 0 = V_1 ((i_{d1} - i_{d2}) \cos \theta_1 + (i_{q1} - i_{q2}) \sin \theta_1) \\ - V_1 \sum_{j=1}^n V_j Y_{1j} \cos(\theta_1 - \theta_j - \phi_{1j}) \end{aligned} \quad (6)$$

$$\begin{aligned} 0 = V_1 ((i_{d1} - i_{d2}) \sin \theta_1 - (i_{q1} - i_{q2}) \cos \theta_1) \\ - V_1 \sum_{j=1}^n V_j Y_{1j} \sin(\theta_1 - \theta_j - \phi_{1j}) \end{aligned} \quad (7)$$

and at bus 2

$$\begin{aligned} 0 = V_2 (i_{d2} \cos \theta_2 + i_{q2} \sin \theta_2) \\ - V_2 \sum_{j=1}^n V_j Y_{2j} \cos(\theta_2 - \theta_j - \phi_{2j}) \end{aligned} \quad (8)$$

$$\begin{aligned} 0 = V_2 (i_{d2} \sin \theta_2 - i_{q2} \cos \theta_2) \\ - V_2 \sum_{j=1}^n V_j Y_{2j} \sin(\theta_2 - \theta_j - \phi_{2j}). \end{aligned} \quad (9)$$

From the figure, the following quantities are defined:

- P_1, Q_1 powers injected from the system into bus 1;
- P_2, Q_2 powers injected from bus 2 into the system;
- P_{sh}, Q_{sh} powers injected by the shunt converter;
- P_{se}, Q_{se} powers injected by the series converter;

P_{dc} active power from shunt to series converter less losses;

P_{inj}, Q_{inj} powers injected by the series transformer
 $= V_{inj} I_2^*$ where $V_{inj} = V_1 - V_2$.

III. NEW UPFC CONTROL

A. Series Control

The control objective for the series portion of the UPFC is inject a variable series voltage into the line such that the line powers track a desired active line power P_2^* and desired reactive power Q_2^* . The target values are chosen for the particular power system application and may be chosen to be a constant value or to damp oscillations. In addition, the shunt portion of the UPFC is utilized to maintain the bus voltage at the sending end (bus 1 in Fig. 2) and the dc link capacitor voltage. The inputs α_1 , k_1 , α_2 , and k_2 are controlled to achieve these objectives.

Starting with the series portion of the UPFC, the desired powers are converted into desired currents i_{d2}^* and i_{q2}^* through

$$\begin{bmatrix} i_{d2}^* \\ i_{q2}^* \end{bmatrix} = \begin{bmatrix} \cos \theta_2 & \sin \theta_2 \\ \sin \theta_2 & -\cos \theta_2 \end{bmatrix}^{-1} \begin{bmatrix} P_2^*/V_2 \\ Q_2^*/V_2 \end{bmatrix}. \quad (10)$$

Note that in per unit, the current in both windings of the series transformer are the same value, therefore the desired line power flows are used to calculate the target series currents.

To track the target, new state variables e_{d2} and e_{q2} are defined such that

$$e_{d2} = i_{d2}^* - i_{d2} \quad (11)$$

$$e_{q2} = i_{q2}^* - i_{q2} \quad (12)$$

leading to new state equations

$$\begin{aligned} \frac{d}{dt}e_{d2} &= \frac{d}{dt}i_{d2}^* + \frac{\omega_s R_{s2}}{L_{s2}}i_{d2} \\ &\quad - \omega i_{q2} - \frac{\omega_s k_2}{L_{s2}} \cos(\alpha_2 + \theta_1)V_{dc} \\ &\quad + \frac{\omega_s}{L_{s2}}(V_2 \cos \theta_2 - V_1 \cos \theta_1) \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{d}{dt}e_{q2} &= \frac{d}{dt}i_{q2}^* + \frac{\omega_s R_{s2}}{L_{s2}}i_{q2} \\ &\quad + \omega i_{d2} - \frac{\omega_s k_2}{L_{s2}} \sin(\alpha_2 + \theta_1)V_{dc} \\ &\quad + \frac{\omega_s}{L_{s2}}(V_2 \sin \theta_2 - V_1 \sin \theta_1). \end{aligned} \quad (14)$$

Let control inputs be defined as

$$u_{21} = k_2 \cos \alpha_2 \quad (15)$$

$$u_{22} = k_2 \sin \alpha_2. \quad (16)$$

A positive definite Lyapunov function is given by

$$V = \frac{c}{2}e_{d2}^2 + \frac{c}{2}e_{q2}^2, c > 0. \quad (17)$$

The derivative of V is given by

$$\dot{V} = p_1 u_{21} + p_2 u_{22} + p_3 - c \frac{R_{s2} \omega_s}{L_{s2}} (e_{d2}^2 + e_{q2}^2) \quad (18)$$

where

$$\begin{aligned} p_1 &= -c \frac{\omega_s}{L_{s2}} V_{dc} (e_{d2} \cos \theta_1 + e_{q2} \sin \theta_1) \\ p_2 &= c \frac{\omega_s}{L_{s2}} V_{dc} (e_{d2} \sin \theta_1 - e_{q2} \cos \theta_1) \\ p_3 &= c \left(e_{d2} \frac{d}{dt} i_{d2}^* + e_{q2} \frac{d}{dt} i_{q2}^* \right) \\ &\quad + c \frac{R_{s2} \omega_s}{L_{s2}} (e_{d2} i_{d2}^* + e_{q2} i_{q2}^*) \\ &\quad - c \omega (e_{d2} i_{q2}^* - e_{q2} i_{d2}^*) \\ &\quad + c \frac{\omega_s}{L_{s2}} (e_{q2} (V_2 \sin \theta_2 - V_1 \sin \theta_1) \\ &\quad + e_{d2} (V_2 \cos \theta_2 - V_1 \cos \theta_1)). \end{aligned}$$

The derivative \dot{V} is guaranteed to be negative if

$$p_1 u_{21} + p_2 u_{22} + p_3 = -c_2 (e_{d2}^2 + e_{q2}^2), c_2 > 0. \quad (19)$$

Therefore from Lyapunov's second theorem on stability [9], this system is asymptotically stable if (19) is satisfied provided the control inputs are selected as

$$\begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = C_2^{-1} \left\{ E_2 - \begin{bmatrix} \omega_s \frac{R_{s2}}{L_{s2}} - c_2 & -\omega_s \\ \omega_s & \omega_s \frac{R_{s2}}{L_{s2}} - c_2 \end{bmatrix} \begin{bmatrix} e_{d2} \\ e_{q2} \end{bmatrix} \right\} \quad (20)$$

where

$$\begin{aligned} C_2 &= \frac{\omega_s V_{dc}}{L_{s2}} \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix} \\ E_2 &= \begin{bmatrix} \frac{\omega_s}{L_{s2}} (R_{s2} i_{d2} + V_2 \cos \theta_2 - V_1 \cos \theta_1) - \omega i_{q2} + \frac{d}{dt} i_{d2}^* \\ \frac{\omega_s}{L_{s2}} (R_{s2} i_{q2} + V_2 \sin \theta_2 - V_1 \sin \theta_1) + \omega i_{d2} + \frac{d}{dt} i_{q2}^* \end{bmatrix}. \end{aligned}$$

Equations (15) and (16) can be solved for k_2 and α_2 from

$$k_2 = \sqrt{u_{21}^2 + u_{22}^2} \quad (21)$$

and

$$\alpha_2 = \begin{cases} \tan^{-1} \frac{u_{22}}{u_{21}} & u_{21} > 0 \\ \tan^{-1} \frac{u_{22}}{u_{21}} + \pi & u_{21} < 0 \\ \sin^{-1} \frac{u_{22}}{k_2} & u_{21} = 0. \end{cases} \quad (22)$$

Both k_2 and α_2 are limited to bound the magnitude of the injected current and therefore limit the injected active and reactive powers. The parameter c_2 is usually chosen to be a large number and may be tuned to obtain the desired damping time frame.

A finer control may be achieved however, by introducing an additional parameter such that (20) becomes

$$C_2 \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = E_2 - \begin{bmatrix} \omega_s \frac{R_{s2}}{L_{s2}} - c_2 & -\omega_s \\ \omega_s & \omega_s \frac{R_{s2}}{L_{s2}} - c_2 \end{bmatrix} \begin{bmatrix} e_{d2} \\ e_{q2} \end{bmatrix}. \quad (23)$$

By splitting c_2 into two separate parameters (c_{d2} and c_{q2}), a weighted control is possible whereby i_{d2} or i_{q2} can be more tightly controlled, and subsequently P_2 or Q_2 .

B. Shunt Control

The control objective for the shunt portion of the UPFC is two-fold. The first objective is to regulate the shunt bus voltage magnitude at the reference value. This is similar to the voltage control aspect of a STATCOM. The second objective is to maintain the dc link capacitor voltage. The control of many voltage sourced converters (such as the STATCOM and SSSC) requires a near constant dc link voltage to have effective control. While the dc link capacitor may discharge briefly during transients to provide active power to the system, a significant voltage sag may result in severe control degradation as the converter is no longer able to inject the desired current into the system. The voltage magnitude at the shunt bus is primarily impacted by injected reactive power, whereas the dc link voltage is primarily impacted by the active power absorbed by the shunt converter to charge the capacitor. Therefore a similar control approach for the STATCOM can be derived such that

$$\begin{bmatrix} i_{d2}^* - i_{d1}^* \\ i_{q2}^* - i_{q1}^* \end{bmatrix} = \begin{bmatrix} \cos \theta_1 & \sin \theta_1 \\ \sin \theta_1 & -\cos \theta_1 \end{bmatrix}^{-1} \begin{bmatrix} P_1^*/V_1 \\ Q_1^*/V_1 \end{bmatrix}. \quad (24)$$

The desired reactive power Q_{sh}^* is the reactive power required for voltage support at the shunt bus and may be chosen independently. The desired active power P_{sh}^* however is not independent of P_{se}^* . The shunt active power must account for the losses in the shunt and series branches of the UPFC and is therefore

$$P_1^* = -(P_{se}^* + P_{loss}) \quad (25)$$

where

$$P_{loss} = R_{s2} (i_{d2}^2 + i_{q2}^2) + R_{s1} (i_{d1}^2 + i_{q1}^2) + \frac{V_{dc}^2}{R_{dc}}. \quad (26)$$

It is not possible to *a priori* designate an exact target value for P_{loss} , but a target value can be estimated. From the series control, values for i_{d2}^* and i_{q2}^* can be obtained. One objective of the shunt portion is to maintain V_{dc} at a constant value, thus V_{dc}^* can also be specified. Thus

$$\begin{aligned} P_{loss}^* &= R_{s2} \left((i_{d2}^*)^2 + (i_{q2}^*)^2 \right) \\ &+ R_{s1} \left((i_{d1}^*)^2 + (i_{q1}^*)^2 \right) \\ &+ \frac{(V_{dc}^*)^2}{R_{dc}} + k(V_{dc}^* - V_{dc}). \end{aligned} \quad (27)$$

The last term in (27) reflects the dependence of P_{loss} on V_{dc} . The constant k is a nonnegative number. Substituting these values into (24) yields

$$\begin{aligned} & \cos \theta_1 (i_{d_2}^* - i_{d_1}^*) + \sin \theta_1 (i_{q_2}^* - i_{q_1}^*) \\ &= \left(P_{se}^* + R_{s_1} \left((i_{d_1}^*)^2 + (i_{q_1}^*)^2 \right) \right. \\ & \quad \left. + R_{s_2} \left((i_{d_2}^*)^2 + (i_{q_2}^*)^2 \right) \right. \\ & \quad \left. + \frac{(V_{dc}^*)^2}{R_{dc}} + k (V_{dc}^* - V_{dc}) \right) / V_1 \end{aligned} \tag{28}$$

$$\sin \theta_1 (i_{d_2}^* - i_{d_1}^*) - \cos \theta_1 (i_{q_2}^* - i_{q_1}^*) = Q_1^* / V_1. \tag{29}$$

These equations can then be solved for $i_{d_1}^*$ and $i_{q_1}^*$. Since these equations are nonlinear, to have an exact solution for $i_{d_1}^*$ and $i_{q_1}^*$ requires the iterative solution of (28) and (29). In most cases however, a fairly close approximation can be obtained by solving

$$\begin{bmatrix} i_{d_1}^* \\ i_{q_1}^* \end{bmatrix} = -\frac{1}{V_1} \begin{bmatrix} \cos \theta_1 + 2i_{d_1}^0 \frac{R_{s_1}}{V_1} & \sin \theta_1 + 2i_{q_1}^0 \frac{R_{s_1}}{V_1} \\ \sin \theta_1 & -\cos \theta_1 \end{bmatrix}^{-1} \\ \times \begin{bmatrix} P_2^* + R_{s_2} \left((i_{d_2}^*)^2 + (i_{q_2}^*)^2 \right) + \frac{(V_{dc}^*)^2}{R_{dc}} + k(V_{dc}^* - V_{dc}) \\ Q_1^* \end{bmatrix}$$

where $i_{d_1}^0$ and $i_{q_1}^0$ are the initial shunt current values.

Following the same procedure above, new state variables e_{d_1} and e_{q_1} are defined

$$e_{d_1} = i_{d_1}^* - i_{d_1} \tag{30}$$

$$e_{q_1} = i_{q_1}^* - i_{q_1} \tag{31}$$

and

$$\begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = C_1^{-1} \left\{ E_1 - \begin{bmatrix} \omega_s \frac{R_{s_1}}{L_{s_1}} - c_1 & -\omega_s \\ \omega_s & \omega_s \frac{R_{s_1}}{L_{s_1}} - c_1 \end{bmatrix} \begin{bmatrix} e_{d_1} \\ e_{q_1} \end{bmatrix} \right\} \tag{32}$$

where

$$C_1 = \frac{\omega_s V_{dc}}{L_{s_1}} \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix}$$

$$E_1 = \begin{bmatrix} \frac{\omega_s}{L_{s_1}} (R_{s_1} i_{d_1}^* + V_1 \cos \theta_1) - \omega i_{q_1}^* + \frac{d}{dt} i_{d_1}^* \\ \frac{\omega_s}{L_{s_1}} (R_{s_1} i_{q_1}^* + V_1 \sin \theta_1) + \omega i_{d_1}^* + \frac{d}{dt} i_{q_1}^* \end{bmatrix}$$

and

$$u_{11} = k_1 \cos \alpha_1$$

$$u_{12} = k_1 \sin \alpha_1$$

from which k_1 and α_1 can be determined similar to (21) and (22).

IV. CONTROL IMPLEMENTATION

The series and shunt controls proposed in the previous section are first applied to the small test system shown in Fig. 3. This system is a two-area system with one low frequency inter area mode in the 1.0-Hz range. Bus 11 is a virtual bus added.

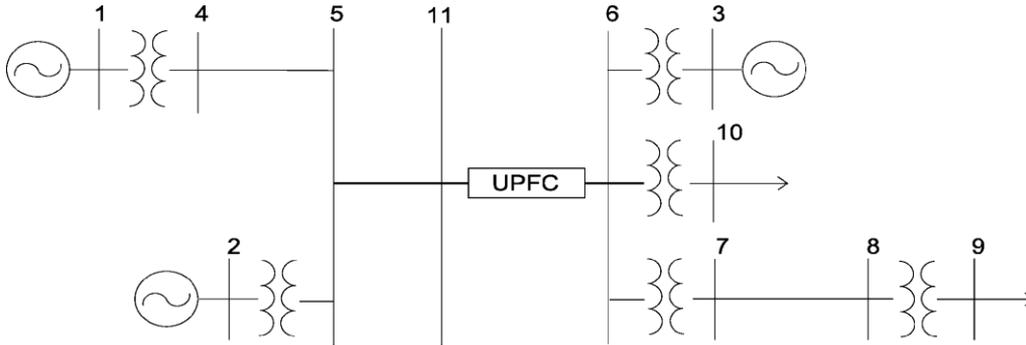


Fig. 3. Two-area test system.

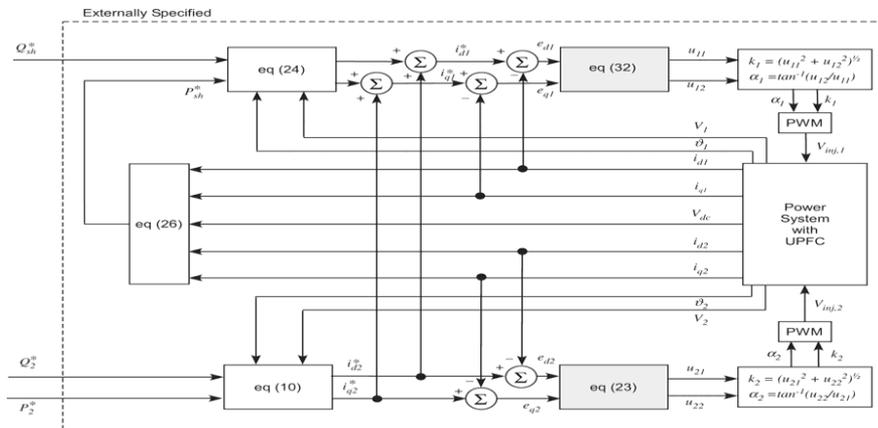


Fig. 4. Control framework.

System the generators are modeled as two-axis generators with a simple dc exciter, voltage regulator, and turbine/governor. The generator and network equations are given in the Appendix. The primary objective of the UPFC is to damp the resulting inter area oscillations. Additional objectives are to maintain the dc link capacitor voltage V_{dc} and the voltage at bus 6. The pro- posed controller is compared against a PI controller where

$$\begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = \begin{bmatrix} K_{1dd} & K_{1dq} \\ K_{1qd} & K_{1qq} \end{bmatrix} \begin{bmatrix} e_{d1} \\ e_{q1} \end{bmatrix} + \begin{bmatrix} K_{1ddI} & K_{1dqI} \\ K_{1qdI} & K_{1qqI} \end{bmatrix} \int \begin{bmatrix} e_{d1} \\ e_{q1} \end{bmatrix} \quad (33)$$

$$\begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = \begin{bmatrix} K_{2dd} & K_{2dq} \\ K_{2qd} & K_{2qq} \end{bmatrix} \begin{bmatrix} e_{d2} \\ e_{q2} \end{bmatrix} + \begin{bmatrix} K_{2ddI} & K_{2dqI} \\ K_{2qdI} & K_{2qqI} \end{bmatrix} \int \begin{bmatrix} e_{d2} \\ e_{q2} \end{bmatrix} \quad (34)$$

The PI parameters were initially chosen via standard procedures 10 and then further refined using a genetic algorithm to produce the best results possible 11. For an even comparison, both the PI controller and the proposed controller are based on the same control scheme shown in Fig. 4. This control is based on dq series and shunt current injections as determined by (10) and (24). Fig. 4 shows the control diagram for the controllers. The only difference between the proposed controller and the PI controller is the shaded blocks, in which (23) is replaced by (33) and (32) is replaced by (34). However, note that the PI controller requires 16 parameters whereas the proposed controller has three parameters (c_1, c_2, k). The parameters for the

proposed control are chosen to be large positive constants and require very little tuning. The parameters for both controllers are given in Tables I and II for Cases I and II.

TABLE I
PROPOSED CONTROLLER PARAMETERS

c_d	1000
c_q	1000
k	10

TABLE II
PI CONTROLLER PARAMETERS FOR CASES I AND II

K_{1dd}	-6.86e-03	K_{2dd}	1.21e-01
K_{1dq}	-2.98e-01	K_{2dq}	-7.72e-05
K_{1qd}	5.43e-02	K_{2qd}	-3.51e-03
K_{1qq}	3.89e-02	K_{2qq}	-4.16e-01
K_{1ddI}	2.34e-01	K_{2ddI}	-9.67e-02
K_{1dqI}	1.26e-02	K_{2dqI}	-9.28e-04
K_{1qdI}	2.59e-02	K_{2qdI}	-1.46e-01
K_{1qqI}	1.96e-02	K_{2qqI}	-8.00e-00

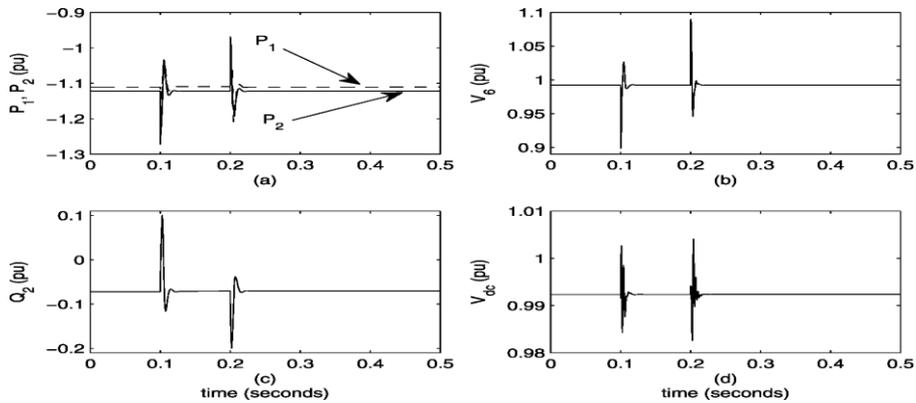


Fig. 5. UPFC control values—Case I.

It is possible to neglect the cross-coupling terms in the PI controller (K_{1dq} , K_{1qd} , K_{2dq} , etc.) to yield only eight parameters that need to be tuned. There is a slight degradation in performance when these mutual affect terms are neglected.

A. Case I

At 0.1 s, a three-phase-to-ground fault is applied to the system on bus 10. The fault is cleared at 0.2 s without removing any lines. The UPFC is designed to control four quantities to a reference value: P_2 , Q_2 , V_1 and V_{dc} . These results are summarized in Fig. 5 for constant reference values. The active power P_1 is also shown in Fig. 5(a). Note that since P_2 and Q_2 are constant, bus 5 simulates a constant ‘‘PQ’’ bus, thus the fault on bus 10 does not sufficiently propagate into area 1. Since the active and reactive power at bus 5 only varies slightly, generators 1 and 2 only experience slight variations due to intra-area oscillations, but the interarea oscillations are substantially reduced. These are shown (along with ω_3) in Fig. 6. Similarly, bus 6 is essentially a ‘‘PV’’ bus with constant power (P_1) and voltage (V_6) and area 1 and area 2 are essentially decoupled by the UPFC.

The system dynamics responses compared against the PI controller and no control are shown in Figs. 7–10 for the parameters given in Tables I and II. In Figs. 7–9 the dotted lines indicate the uncontrolled system dynamic responses. The system is stable but exhibits sustained interarea oscillations. The response of the system with the PI controller is shown with the thin line. The proposed controller response is the bold line. Fig. 7 shows the active power flow on the tie line between areas 1 and 2. Both controllers exhibit good damping response as compared to the

GUO *et al.*: AN IMPROVED UPFC CONTROL FOR OSCILLATION DAMPING

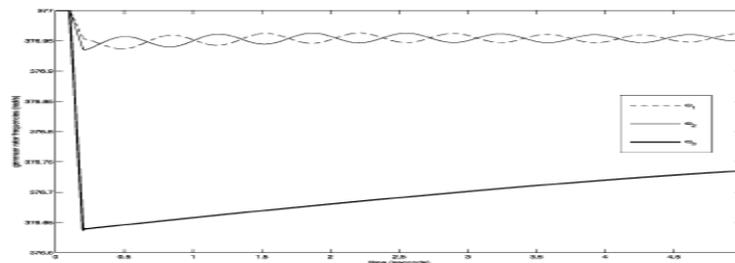


Fig. 6. Generator frequencies (rad/s)—proposed control—Case I.

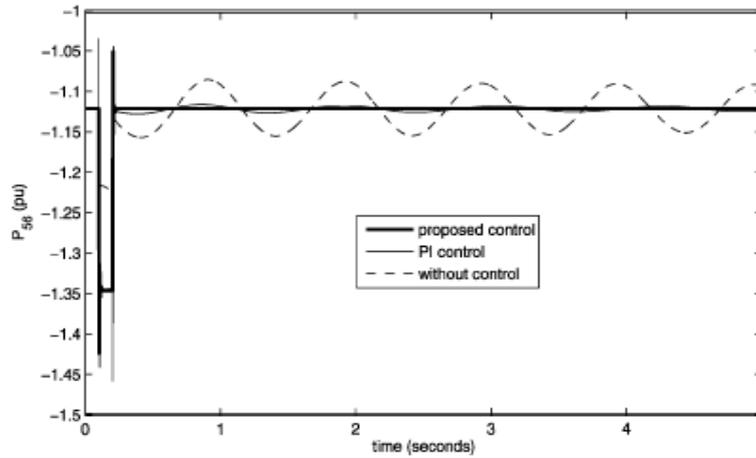


Fig. 7. Active power between buses 5 and 6—Case I.

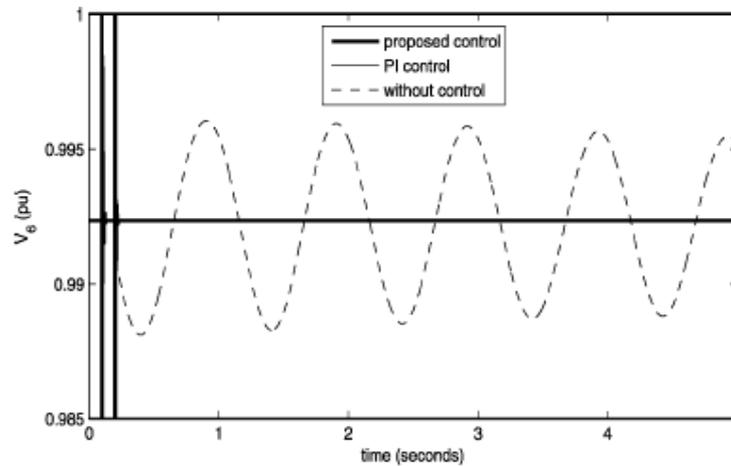


Fig. 8. Voltage magnitude at bus 6—Case I.

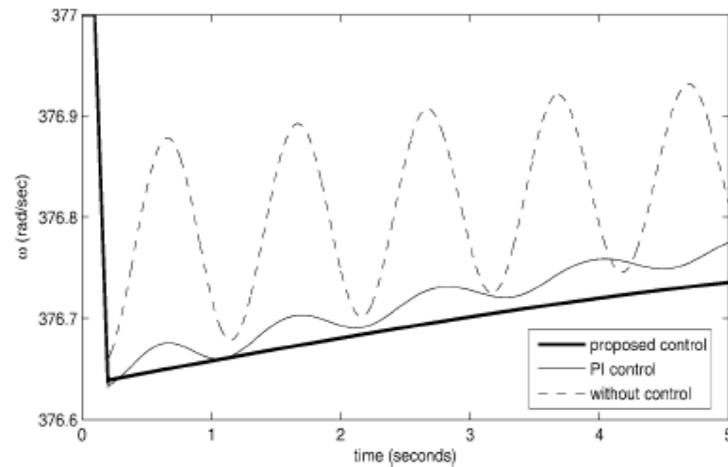


Fig. 9. Generator frequencies—Case I.

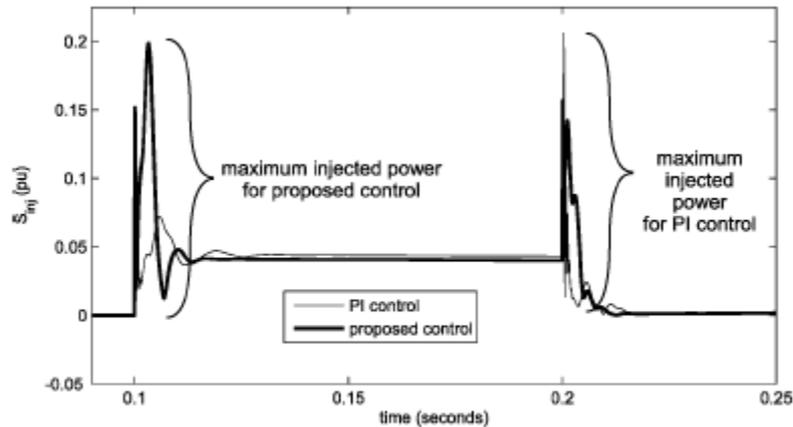


Fig. 10. Injected UPFC power for PI and proposed control—Case I.

UN damped case. Fig. 8 shows the voltage at bus 6. The results of the PI and proposed controller are virtually indistinguishable and both show excellent voltage control. Fig. 9 shows the frequency dynamics of generator 3. Both controllers show similar damping as compared to the uncontrolled case. Fig. 10 shows the injected power during the fault which is less than 20% of the steady-state power flow on the line. Note that the power injection for the proposed control is greatest at the fault initiation (0.1 s), whereas the PI control is maximum at the fault clearing (0.2 s). During the fault (after transients), the power injection of both controllers is nearly the same. This similar behavior allows an even comparison between controllers. Both controllers exhibit a “spike” at fault initiation and clearing. This is due to the slight time delay before the controller reacts. This is an artifact of the transmission model used—the algebraic network equations provide instantaneous propagation of the fault behavior throughout the system. In actuality, the fault propagation rate depends on the dynamics of the RLC transmission network and is not instantaneous—therefore in practical situations, the spike magnitude would probably not be as high.

The duration of the spike depends on several factors: the controller design, the sampling rate of the digital signal processor, the switching frequency of the power electronics, and the microprocessor speed. In the results shown, the only factor apparent is the controller design and the “sampling rate” which can be considered to be the time step length of the time domain simulation. In this case, both controllers can be said to act equally fast.

B. Case II

One substantial difference between linear and nonlinear controllers is that linearized control is typically optimized around a single operating point, but a nonlinear controller is relatively independent of a particular operating point. To illustrate the differences in controller response, the same controllers (same parameters) are now subjected different operating conditions. In this case, the same system is used except that the generator inertias are altered slightly to change the frequency of the interarea mode. The power flow response to the same disturbance is shown in Fig. 11. The uncontrolled system is stable with sustained oscillations. The proposed controller still exhibits good damping characteristics with very rapid damping. The PI controller still damps the oscillations but at a much slower rate. To be effective, the PI controller must be retuned to the new frequency of oscillation.

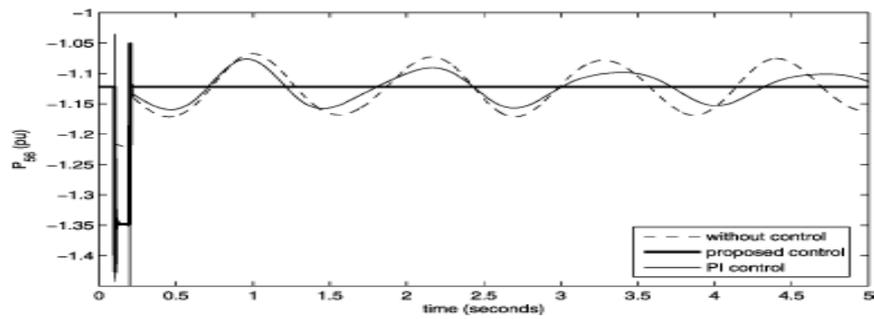


Fig. 11. Active power between buses 5 and 6—Case II.

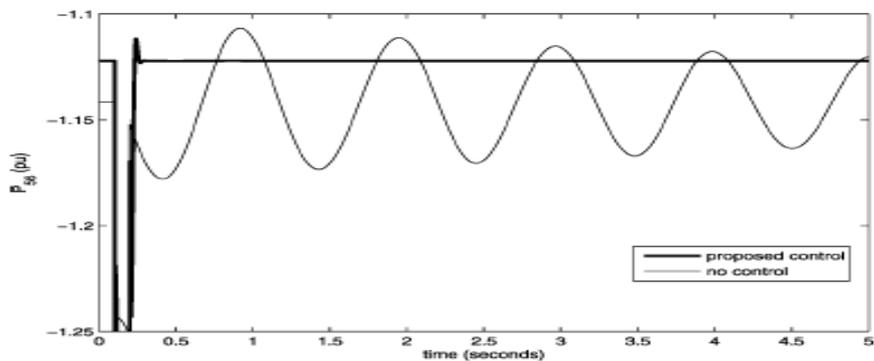


Fig. 12. Active power between buses 5 and 6—Case III.

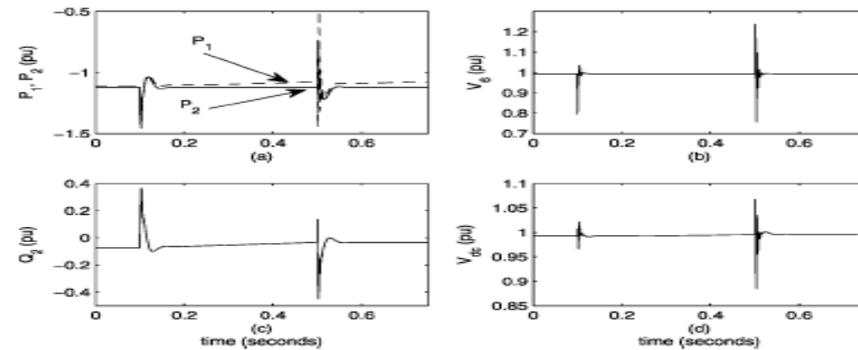


Fig. 13. UPFC control values—Case III.

Power exceeds its limits. In this case, the fault is not removed until 0.5 s. The resulting active power flow on line 5–6 is shown in Fig. 12 has roughly twice the swing magnitude of Case I. Similar to Fig. 5, all of the controlled parameters are shown in Fig. 13. However, the injected power exceeds 20% of the steady-state power flow on the line (assumed to be the UPFC limit). Therefore, constraints are put on the injected power to keep it below the UPFC limit. The injected power under limited and non-limited operation is shown in Fig. 14. The dashed line indicates the maximum injected power limit. The resulting active power flows on line 5–6 are shown in Fig. 15. Note that although the performance of the UPFC in maintaining a constant active power on the line is somewhat degraded, it rapidly brings the power flow back to the reference once the fault is cleared

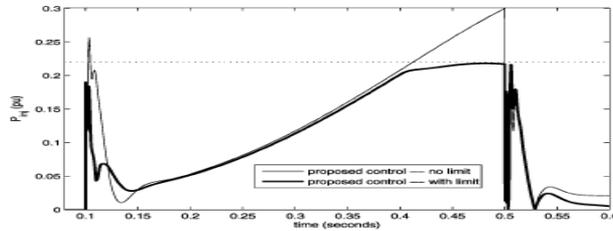


Fig. 14. UPFC injected power—Case III.

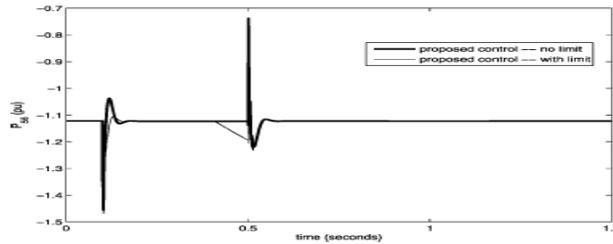


Fig. 15. Active power between buses 5 and 6—Case III.

D. Case IV

Next, the proposed controller’s effectiveness was tested in the 118 bus test system shown in Fig. 16. The UPFC was placed on line 26–30 and a three-phase-to-ground fault occurred on bus 43 at 0.2 s and was cleared at 0.4 s. The UPFC objective was to regulate the post fault active power flow on the line. Without the controller, highly nonlinear active power oscillations are induced by the fault as shown in Fig. 17. As in the small test system, the proposed controller exhibits near immediate control. Fig. 18 shows the injected power from the UPFC required to achieve the damping shown in Fig. 17. Note that the amount of injected active power is considerably less than the amount of active power flow on the line and is only a few percent of the line rating.

E. Case V

In the last case, two UPFCs are placed in the 118 bus system. They are placed on lines 5–11 (at bus 5) and 7–12 (at bus 7). These placements were chosen such that the two UPFCs were in close proximity to see if they adversely affected each other’s performance. In this case, the reference values of each UPFC’s powers were altered such that (SS indicates steady-state): These values were chosen to show the impact of both positive and negative reference values on the behavior of the UPFCs. The dynamic impact of these changes are shown in Figs. 19–21. Fig. 19 shows all of the series and shunt powers of the two UPFCs in response to the commanded reference changes. The top plot shows the active series powers of the UPFCs. UPFC1 is shown as the solid trace whereas UPFC2 is the dashed trace. The middle plot shows the series reactive powers and the bottom plot

GUO et al.: AN IMPROVED UPFC CONTROL FOR OSCILLATION DAMPING

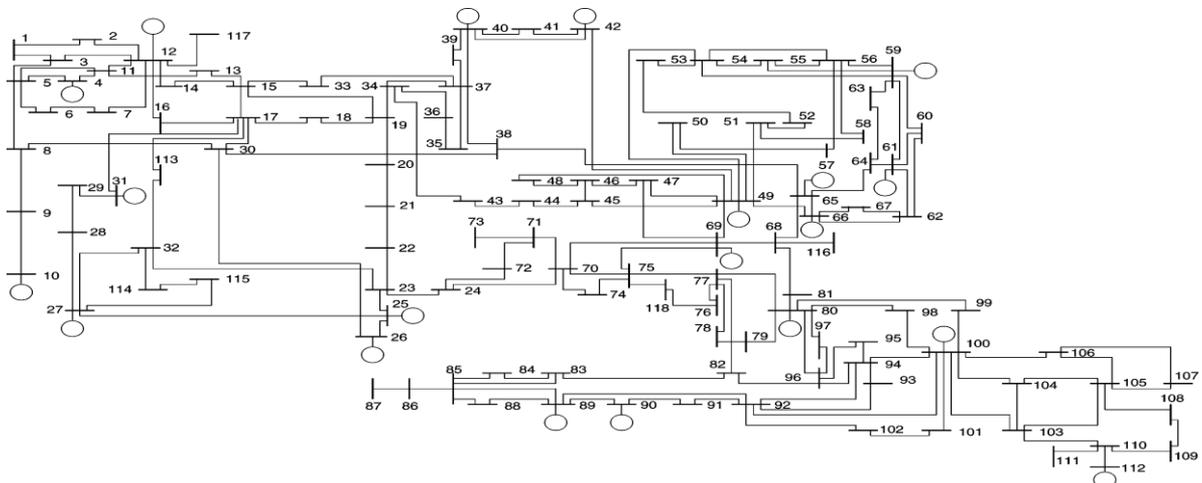


Fig. 16. IEEE 118-bus test system.

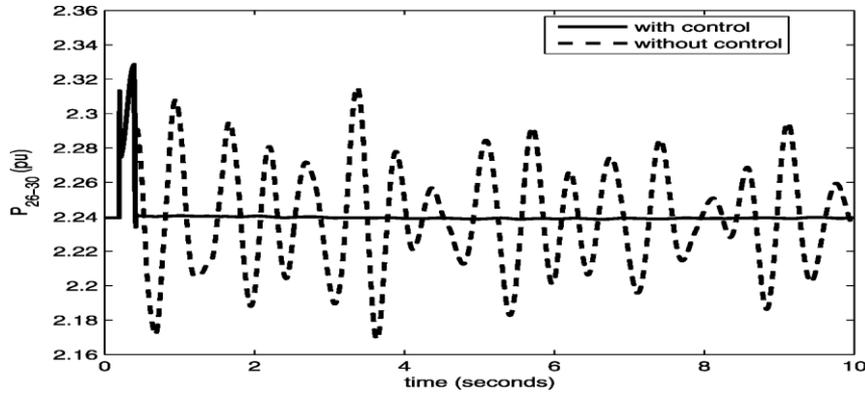


Fig. 17. Active power between buses 26 and 30—Case IV.

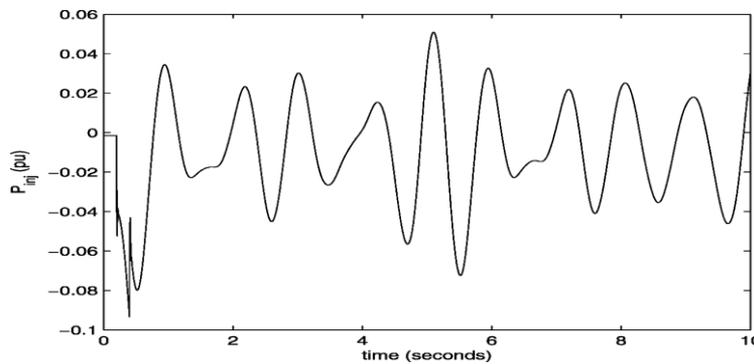


Fig. 18. UPFC injected active power—Case IV.

	UPFC 1 (5-11)			UPFC 2 (7-12)		
	P_2^*	Q_2^*	Q_1^*	P_2^*	Q_2^*	Q_1^*
$0 \leq t \leq 0.1$	SS	SS	SS	SS	SS	SS
$0.1 < t \leq 0.2$	1.5	—	—	-0.3	—	—
$0.2 < t \leq 0.3$	—	-0.3	—	—	0.2	—
$0.3 < t \leq 0.4$	—	—	0.0	—	—	0.0

Shows the shunt injected reactive power. The active power references are commanded to change at 0.1 s. Note that the active powers respond rapidly to the change in commanded powers. There is a slight impact on both the series and shunt reactive

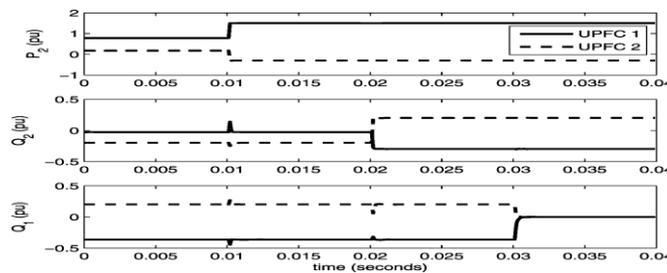


Fig. 19. Active and reactive UPFC powers—Case V.

Powers, but they rapidly return to their commanded values. The series reactive power is commanded to change at 0.2 s. As expected, this change has more impact on the shunt injected reactive power than on the series active power, but once again, the shunt reactive power returns to the commanded value rapidly. Lastly, the shunt reactive power is commanded to change at 0.3 s. This change has little impact on the series active and re- active powers. Fig. 20 shows the dc link capacitor voltage of the two UPFCs throughout the changes. Note that aside from brief small transients at the time of reference changes, the dc link capacitor volt- ages remain constant. This indicates good stability of the control and ensures good performance of

the UPFC. Fig. 21 shows the bus voltage magnitudes at either end of the transmission line on which the two UPFCs are placed. The voltages behave as expected in response to the changes in the active and reactive powers. These results indicate that the proposed control enables independent control of the three independent UPFC attributes (series active power, series reactive power, and shunt reactive power). The shunt active power depends on the charge and discharge of the dc link capacitor and the UPFC losses.

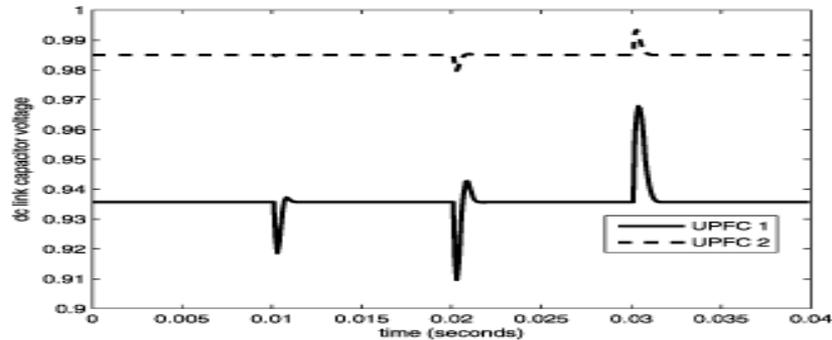


Fig. 20. DC link capacitor voltages—Case V.

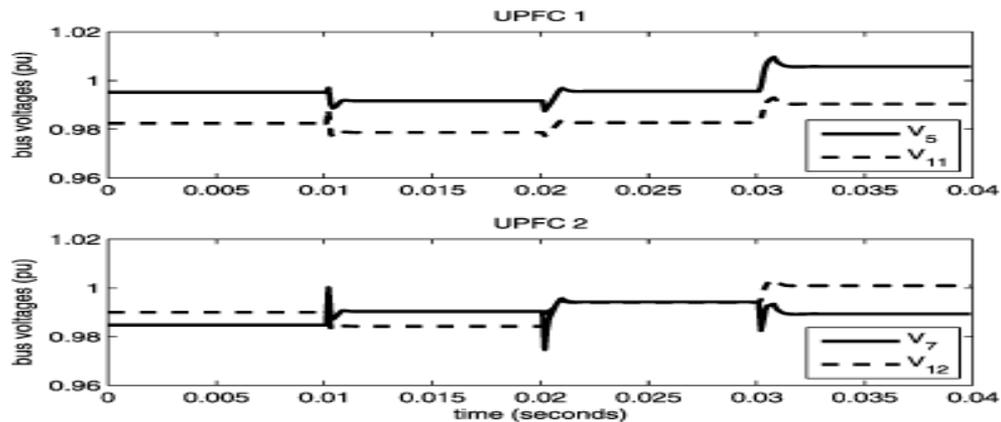


Fig. 21. Bus voltages—Case V.

V. CONCLUSION

In this paper, a new control for the UPFC was proposed. The proposed control exhibited very good performance in damping active power oscillations and maintaining the UPFC shunt bus voltage. It exhibited favorable performance when compared with a PI controller under several operating conditions. The proposed control works well in both large and small systems with rapid dynamic response and independent control. The primary advantages of the proposed control are 1) it works over a wide range of operating conditions, 2) requires only three parameters, and 3) the parameters are easily chosen and do not require considerable tuning effort.

REFERENCES

- [1] B. C. Pal, "Robust damping of inter area oscillations with unified power-flow controller," *Proc. Inst. Elect. Eng. Gen., Transm., Distrib.*, vol. 149, no. 6, pp. 733–738, Nov. 2002.
- [2] B. Chaudhuri, B. C. Pal, A. Zolotas, I. Jaimoukha, and T. Green, "Mixed-sensitivity approach to control of power system oscillations employing multiple FACTS devices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1149–1156, Aug 2003.
- [3] M. M. Farsangi, Y. H. Song, and K. Y. Lee, "Choice of FACTS device control inputs for damping interarea oscillations," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1135–1143, May 2004.
- [4] E. Gholipour and S. Saadate, "Improving of transient stability of power systems using UPFC," *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 1677–1682, Apr. 2005.
- [5] S. Kannan, S. Jayaram, and M. Salama, "Real and reactive power coordination for a unified power flow controller," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1454–1461, Aug. 2004.
- [6] N. Tambey and M. L. Kothari, "Damping of power system oscillations with unified power flow controller (UPFC)," *Proc. Inst. Elect. Eng. Gen., Transm., Distrib.*, vol. 150, no. 2, pp. 129–140, Mar. 2003.
- [7] H. Wang, "A Unified model for the analysis of FACTS devices in damping power system oscillations—Part III: Unified power flow controller," *IEEE Trans. Power Del.*, vol. 15, no. 3, pp. 978–983, Jul. 2000.

Modeling Of Auto Recloser for Smart Grid

Mandar P. Katti, Jangamshetti S. H, Ajay Rege

Abstract: Power Distribution Networks have been operated in an easy and simple unidirectional way. Therefore, no automatism even remote control technology was applied to Ring Main Units or Sectionalizer close to the load in the distribution network. However, increase in load demand has led to instability in system which ultimately resulted into outages in power system. This paper presents the SIMULINK modeling of a control circuit of autorecloser which is one of the most important equipment in smart grid. The modeled control circuit is synchronized with circuit breaker in such a way that after occurrence of fault, breaker recloses automatically without any human interference. The other important aspect of smart grid is to make power system self sufficient. The importance of having Decentralized Generation at Medium voltage and Low voltage level is explained. Whenever transient fault occurs in the system at distribution level, autorecloser avoids outage for longer duration. However, if the fault is persistent then autorecloser isolates only affected part in system and avoids outage in other parts of system.

Keywords: DG, MV, LV, HV, Distributed Energy Resources, Present and Future power system.

I. Introduction

Switchgear and control gear are necessary at every switching point in power systems. The switchgear and control gear industry in India is a fully developed industry, producing and supplying a wide variety of switchgear and control gear items needed by industrial and power sectors. Autorecloser is a circuit breaker equipped with a mechanism that can automatically close the breaker after it has been opened due to a fault [2]. Automatic circuit reclosing is extensively applied to overhead line circuits where a high percentage of faults that occur are transient in nature. The Smart Grid is idea of a better electricity delivery infrastructure. Smart Grid implementations will certainly increase the quantity, quality, and use of information available from advanced sensing, computing, and communications hardware and software [1]. The Smart Grid is idea of a better electricity delivery infrastructure. Smart Grid implementations will certainly increase the quantity, quality, and use of information available from advanced sensing, computing, and communications hardware and software [1].

II. Smart grid and Autorecloser

a. Smart grid

The vision of a smarter grid is to make the electric power system more interactive, interoperable, reliable, and robust—"self-healing".

India's electric grid should make three fundamental improvements to the existing grid:

- 1] Advanced metering to reduce AT&C losses that are at an unacceptably high-level presently

- 2] Automation to measure and control the flow of power to/from consumers on a near real-time basis and improve the system reliability
- 3] Moving to a smart grid to intelligently manage loads, congestion and shortfall From the last point it is clear that to have decentralized generation it will help to manage loads and help to overcome shortages in power.

Due to increasing power demand in far-flung areas it is very difficult to satisfy these demand as it is uneconomical to supply power to these areas. Hence, Decentralized Energy Resources (DER) is used in such areas.

The overall problem when integrating DG in existing networks is that distribution systems is a unidirectional system from the central generation downstream to the consumer. The conventional protection systems were designed in common Medium Voltage (MV) and Low Voltage level (LV) distribution networks [3].

b. Autorecloser

Automatic reclosing is widely adopted in medium voltage networks. Automatic reclosing is easy to implement in a radial distribution network. It becomes problematic when distributed generation is introduced to the network [4]. The ARD model integrated to the circuit breaker developed in MATLAB SIMPOWERSYSTEMS is adapted from the study conducted by MM EL-Saadawi in reference [6] with certain modifications in fast curve block.

Autorecloser functions on the principle of Coordination of Inverse Time Overcurrent Relays with Fuses. The duty of protection equipment is to allow overload currents that occur during operation, yet to prevent impermissible loading of lines and equipment [10]. To avoid damages in the case of short-circuits the relevant equipment must be tripped in the shortest possible time [3]. On the other hand only few feeders or loads as possible should be disconnected from supply. The protection relays available in the power system must recognize the fault, perform tripping themselves or give trip commands for the relevant switching device.

III. Autorecloser in present system and future power system

Figure 4.1 shows the control circuit of autorecloser which is located inside the subsystem.

- **Sine wave:** The sine wave block is the representation of AC source that is considered as supply source. For recloser AC (230V) or DC (110V) source can be considered as supply source.
- **RMS (Root Mean Square):** The RMS block is used to measure the root mean square value of the instantaneous current passing through the recloser.
- **Gain:** The gain block is used to obtain peak value of the instantaneous current passing through the recloser.
- **Time-Current Characteristics:** The peak value will pass to two blocks; the first is a Function Block

parameter which contains the fast curve of the recloser (TCC). This fast curve is based upon the IEEE STANDARD INVERSE-TIME characteristic equations. The equation for Time Current Characteristics [3] is given as,

The power system model shown in figure 4.2 is of generation, transmission, distribution system. The focus is basically on distribution system as autorecloser is used in

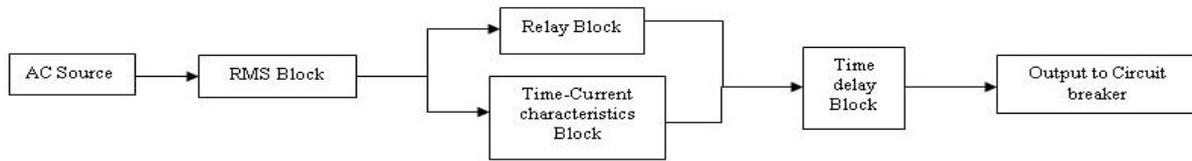


Figure 4.1: Recloser control circuit

$$tt = \frac{\frac{K_d}{\tau_s} TDS}{\left(\frac{I}{I_p}\right)^p - 1} \quad (1)$$

Where,

- tt: Trip time,
- K_d : drag magnet damping factor,
- τ_s : Initial spring torque,
- I: normal current,
- P: constant exponent,
- TDS: Time dial setting,
- I_p : relay pickup

The equation (1) can be further modified as,

$$tt = \frac{A}{M^p - 1} TDS \quad (2)$$

Where,

$$A = \frac{K_d}{\tau_s},$$

$$M = \frac{I}{I_p},$$

The output of this block is a time corresponding to the passing current.

- **Relay:** The next block is a Relay Block which allows its output to switch between two specified values (0, 1). If the current is less than a specific value (reclosers setting) the relay output will stay at zero value, if the current value is greater than that specific value and more the output of the relay will be stick with 1.
- **Variable time delay:** Variable Time Delay block receives the output of the previous two blocks as an input. The output of that block will be either 0 or 1 after a delayed time. If a fault current is passed through the relay; its output signal is 0, and this signal will be delayed (by the variable time delay block) for a short time inversely proportional to the fault current value. The output of the last block is a signal that opens the breaker switch. If the fault is a temporary one, the relay output will be 1, so that the breaker switch closes.
- **Scope:** The signal in the recloser control circuit is monitored at various levels with the help of scope. Basically, there are 4 scopes used at different level in recloser control circuit.

Computational block diagram of present power system

The high percentage of temporary faults allows the application of a device with a dual timing characteristic to coordinate with fuses, Sectionalizer, and other automatic circuit reclosers placed on the system. The number of reclosing events, the reclosing interval delays, fast and time delay (also referred to as slow curve selection, and minimum trip selection must be chosen to satisfy a number of objectives. The block diagram of present power system given in figure 4.2 indicates the presence of autorecloser in the system.

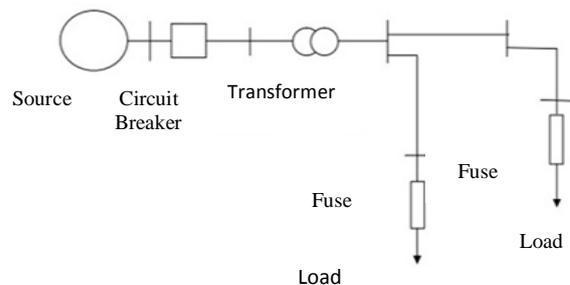


Figure 4.2: Present power system

Computational block diagram of future power system using autorecloser and decentralized generation

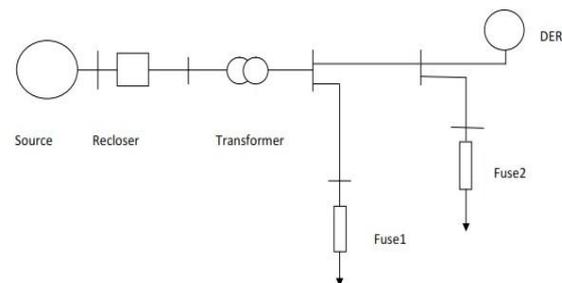


Figure 4.3: Decentralized generation connected to the system.

Figure 4.3 illustrates a Decentralized generation connected to the system. Decentralized generation is defined as it can be defined as the development of small, modular electric

generation close to the point of consumption [11]. DER's impact on the power system industry and these impacts can be categorized as, financial, technical, and regulatory impacts. Integration of DG in distribution networks may impact the network protection system [4].

IV. Results and Discussion

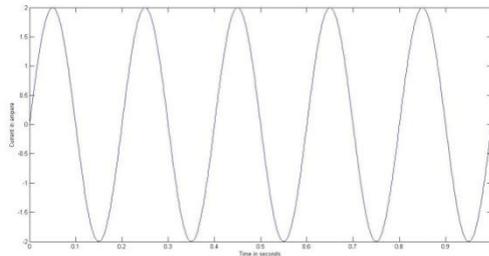


Figure 5.1: Input to recloser control circuit

Figure 5.1 shows the sinusoidal AC instantaneous input fed to the autorecloser. Generally, the autorecloser is fed AC as input or DC in case of power failure. The sinusoidal input is resemblance of AC input fed to the recloser control circuit.

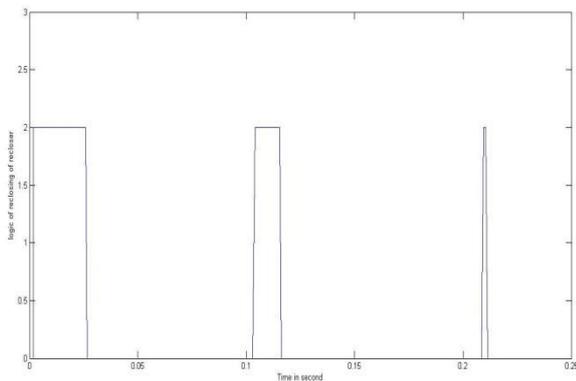


Figure 5.2: Reclosing attempts carried out by recloser

Figure 5.2 shows the reclosing attempt carried out by autorecloser when fault is still persistent even after first reclosing attempt. The frequency of reclosing depends upon system design, its capacity to withstand the continuous making and breaking of the breaker. Ideally, for distribution system the reclosing process is limited to 3 to 4 attempts to avoid major damage to the circuits.

When a temporary fault occurs in the system, the breaker contacts get separated and reclosing is initiated. As the fault is temporary, the recloser generally resumes supply in first attempt itself. However, sometimes it may take 2 or 3 attempts.

When persistent fault occurs in the system, as breakers keeps on tripping due to presence of fault and which results into failure in reclosing of breaker. Hence, system goes into lockout mode.

Various fault conditions in the present power system

The power system block diagram is constructed in SIMULINK environment and is simulated under various fault conditions

Figures 5.3 and 5.4 clearly states that when phase to phase fault occurs in the system as no recloser is present in the system, the power system is needed to be checked for any existence of fault. This leads to increase in the outage time. In present power system, engineers has to go and

manually check whether the fault is existing or not and after confirmation can resume the supply. But, the other parts of the system which are not affected by fault, faces outage.

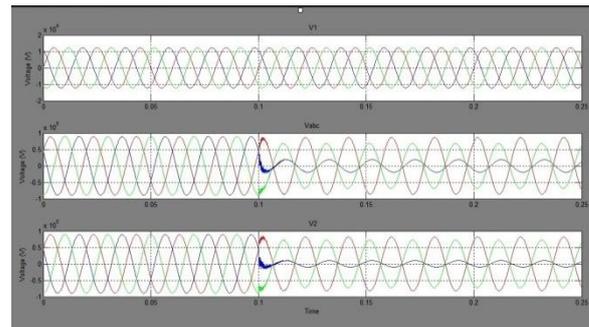


Figure 5.3: Voltage output near generation for during phase to phase fault.

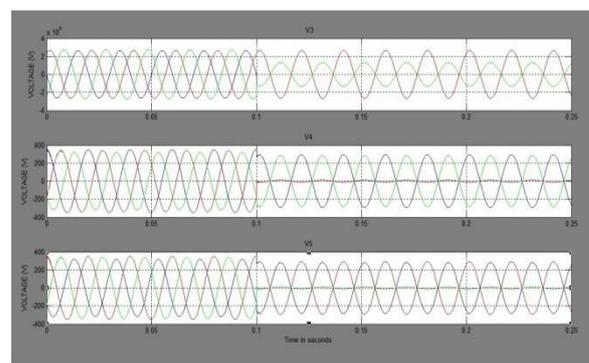


Figure 5.4: Voltage output near load during phase to phase fault

Various fault conditions in the future power system

Future power system basically consists of autorecloser and DG systems as parts of distribution system.

Figure 5.5 and 5.6 shows the effect of having autorecloser in distribution system during phase to phase fault in the system. Due to the presence of autorecloser in the distribution system after two attempt the recloser resumes the supply to the affected system. This is possible only if the fault is temporary in nature.

When DG system is present in the system at MV and LV level the reactive power flowing the system is affected. It also affects the autorecloser by causing recloser-fuse mis-co-ordination.

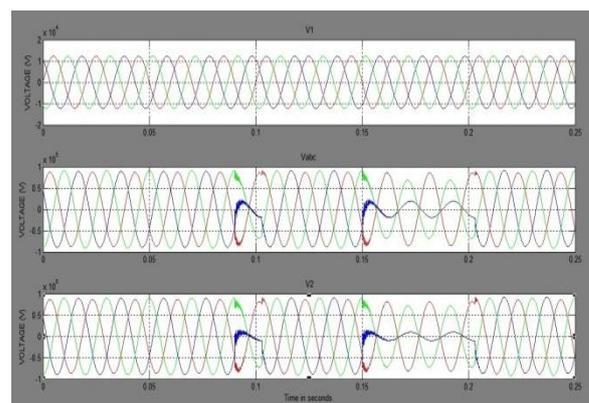


Figure 5.5: Voltage output near generation during phase to phase fault in presence of autorecloser.

Figure 5.7 and 5.8 clearly explains the necessity of converters in DG system. When converters are used it can be observed that reactive power is under control which justifies the presence of converters in DG to obtain balanced output. The outputs shown in figure 5.7 and 5.8 are obtained in absence of autorecloser.

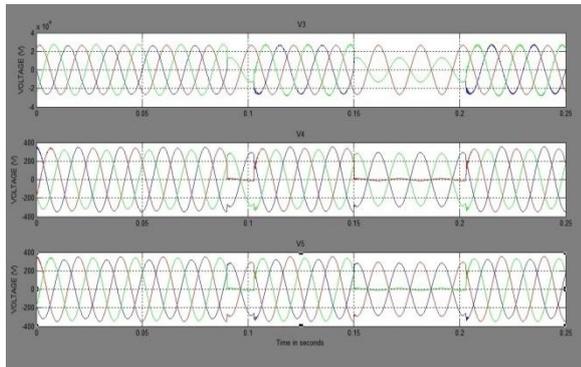


Figure 5.6: Voltage output near load during phase to phase fault in presence of autorecloser

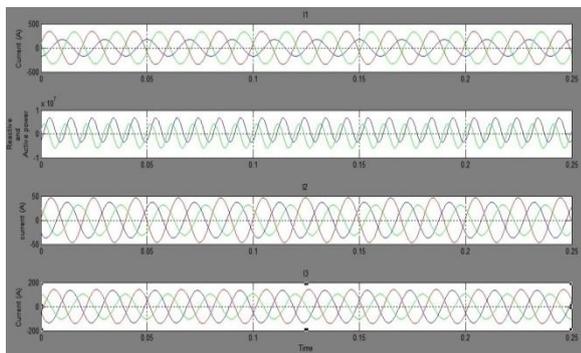


Figure 5.7: Current and reactive power output in presence of DG (15kV) system at distribution level

Figure 5.9 and 5.10 explains that in presence of autorecloser and properly designed the DG the system can overcome any type of temporary faults.

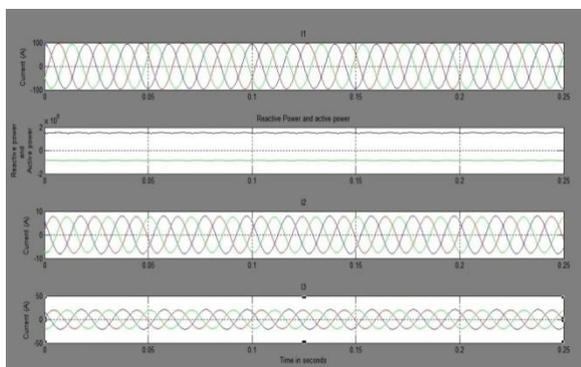


Figure 5.8: Current and reactive power output in presence of DG (415V) system at distribution level.

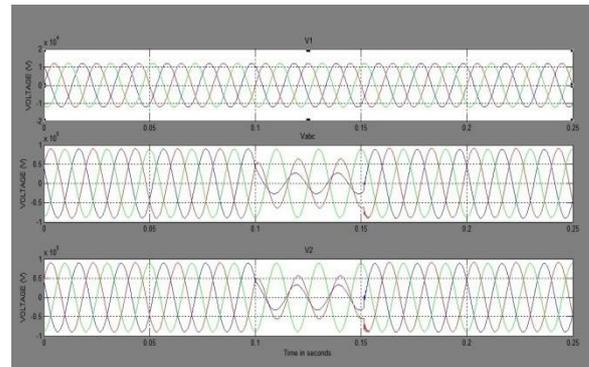


Figure 5.9: Voltage output near generation during phase to phase fault in presence of DG and recloser.

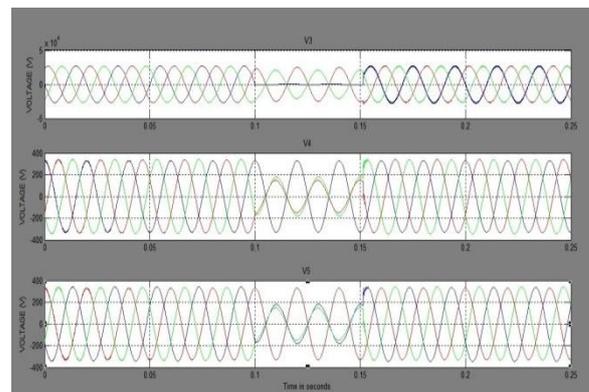


Figure 5.10: Voltage output near load during phase to phase fault in presence of DG and recloser.

Comparison of the results

Table 5.11: Theoretical and simulation values of reclosing time for given values of current.

Current (A)	Theoretical data (Reclosing time in Seconds)	Simulation data (Reclosing time in Seconds)
290	0.5	0.4
320	0.3	0.3
400	0.2	0.2
490	0.15	0.1
550	0.1	0.09
650	0.09	0.07
810	0.05	0.04
1150	0.04	0.03
2000	0.033	0.02
3000	0.025	0.01
4000	0.015	0.009

Table 5.11 illustrates the theoretical and simulated data for given values of the current flowing through the recloser. Figure 5.12 and 5.13 clearly shows the difference in the reclosing time obtained between theoretical data and simulated data.

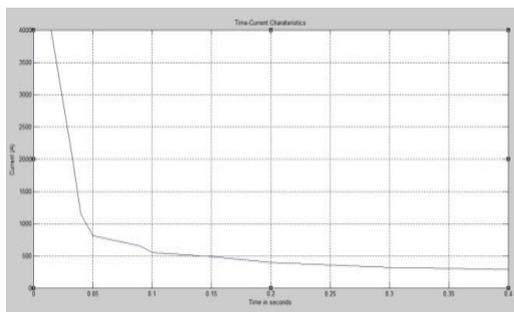


Figure 5.12: Time Current characteristics for theoretical data

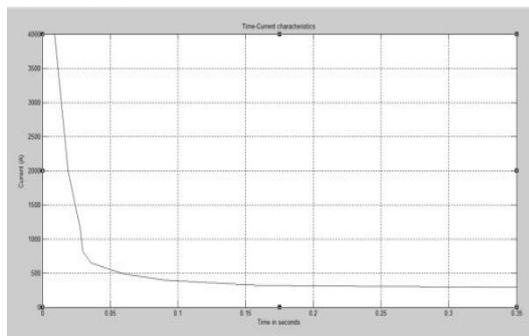


Figure 5.13: Time current characteristics for simulated data.

V. Conclusion

A detailed study is done concerning the application of the autorecloser in the present and future power system which was presented using SIMULINK toolbox. This aimed on improving the voltage, current and reactive power flow in power system. The recloser results obtained are compared to determine the better design of the control circuit. Also, the correct position of the Decentralized Generation (DG) in the distribution system helps in improving system performance during fault.

The results obtained clearly states that the use of autorecloser with some modification will be beneficial if implemented in Indian power system. The use of decentralized generation in the distribution system will help the power system in reducing the outage time by diverting load to unaffected part. However, the decentralized generation system affects the recloser and fuse co-ordination. This problem is solved by selecting a DG system which does not affect recloser fuse co-ordination.

References

- [1] Bruno OPITSCH (Siemens AG – Germany), Manfred HASLINGER (Siemens AG – Austria), Markus SPANGLER (Siemens AG – Germany), “INTELLIGENCE FOR SMART GRIDS LAST MILE”, Paper 0396- SIEMENS LTD.
- [2] J.E Witte, S.R. Mendis, M. T Bishop, and J.A. Kischefskyz, “Computer-Aided Recloser Applications for Distribution Systems”, ISSN-0895-0156/92, 1992 IEEE, July-1992.
- [3] Sukumar M. Brahma, *Student Member, IEEE*, and Adly A. Girgis, *Fellow, IEEE.*, “Development of Adaptive Protection Scheme for Distribution Systems With High Penetration of Distributed Generation”, IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 19, NO. 1, January-2004.

- [4] M.Njozela, *Non-Member, IEEE*, S.Chowdhury, *Member, IEEE*, and S.P.Chowdhury, *Member, IEEE*, “Impacts of DG on the Operation of Auto-Reclosing Devices in a Power Network”, Power and Energy Society General Meeting IEEE, 24-29 July 2011.
- [5] J. Jäger, T. Keil, L. Shang, R. Krebs. , Siemens AG, Germany, “NEW PROTECTION CO-ORDINATION METHODS IN THE PRESENCE OF DISTRIBUTED GENERATION”, 8th International Conference on Developments in Power System Protection Amsterdam (The Netherlands), 5.-8. April 2004.
- [6] M.M. El-Saadawi, “Impact of Distributed Generation on Coordination of Protective Devices”, Dept. of Electrical Engineering, Faculty of Engineering, Mansoura University.
- [7] “SPECIFICATION FOR 11kv OUTDOOR POLE-MOUNTED AUTO-RECLOSER WITH REMOTE COMMUNICATION CAPABILITIES”, UGVCL specifications.
- [8] “Line Protection in Distribution Systems”, Coordination of Inverse Time Overcurrent Relays with Fuses, Siemens PTD EA, Applications for SIPROTEC Protection Relays, 2005.
- [9] Network Protection & Automation Guide – 2011. 219-231.
- [10] Indrapal, Shubhra chaturvedi – “Autorecloser and Sectionalizer”, HANDBOOK OF SWITCHGEARS, Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com) Copyright © 2007 The McGraw-Hill Companies.
- [11] <http://www.isgtf.in/>
- [12] http://www.ibm.com/smarterplanet/in/en/smart_grid/deas/index.html
- [13] http://www.siemens.co.in/en/about_us/index/our_business_segments/energy/power_distribution/smartgrid.htm

Biographies

Mandar P Katti:



Mandar P Katti was born in Ulhasnagar, Maharashtra, India on December, 29th 1987. He obtained his B.E (Electrical) degree from Yadavrao Tasgaonkar Institute of Engineering and Technology, Mumbai, India in 2010 and presently he is persuing M.Tech. (Power and Energy System) in Basaveshwar engineering college, Bagalkot. His areas of interest are Renewable Energy, FACTS, Smart grid.

Dr. Suresh H. Jangamshetti



Dr. Suresh H. Jangamshetti was born in Bijapur, Karnataka, India on May 28, 1963. He obtained his B.E (Electrical) degree from Karnataka University Dharwad in 1985 and Master in Technology (Power Systems) and Ph.D (Wind Energy Systems) from IIT Kharagpur in 1989 and 2000 respectively. His areas of interest include Wind-Solar Energy Systems, Energy Conservation, Computer Applications to Power System and FACTS. He won the "Outstanding IEEE Student Branch Counsellor" award for the year 1996 (R10)

and 2010 (IEEE Bangalore Section) at Basaveshwar Engineering College, Bagalkot, Karnataka, India. He was Fulbright-Nehru Visiting Lecture Fellow at Michigan Technological University, Houghton MI USA during fall 2011. He is working as Professor in the department of Electrical and Electronics at Basaveshwar Engineering College, Bagalkot.

Mr. Ajay Rege

Mr. Ajay Rege was born in Mumbai, India on 10th March, 1962. He obtained B.E. degree in Electrical Engineering in 1984 from Walchand College of Engineering, Sangli then affiliated to Shivaji University, Kolhapur.

He has been associated with Siemens for over 26 years and is presently working at Siemens Ltd. India at Kalwa as General Manager – Marketing & Sales leading the team of Medium Voltage Products Exports.

Unsteady Mixed Convective Heat and Mass Transfer flow through a porous medium in a vertical channel with Soret and Dissipation effects

Sudha Mathew^a, P. Raveendra Nath^b, B. Sreenivasa Reddy^c

^a Research scholar, Department of Mathematics, S.K. University, Anantapur, A.P., India

^b Sri Krishnadevaraya University College of Engineering and Technology, Anantapur, A.P., India.

^c Assistant professor, Department of Mathematics, Yogivemana University, Kadapa

Abstract: Unsteady Hydromagnetic Mixed Convection flow of a viscous, electrically conducting fluid through a porous medium confined in a vertical channel bounded by flat walls. The unsteadiness in the flow is due to the travelling thermal wave is imposed on the bounding walls. The concentration on the walls is maintained constant. A uniform magnetic field of strength H_0 is applied transverse to the boundaries. The coupled equations governing the flow, heat and mass transfer are solved by using the perturbation technique with δ , the aspect ratio as a perturbation parameter. The combined influence of the Soret and dissipation effects on the velocity, temperature, concentration, stress and rate of heat and mass transfer are discussed in detail.

Keywords: CFD, Mixed Convection, Heat Transfer, Mass Transfer, Dissipation

I. Introduction

The time dependent thermal convection flows have applications in chemical engineering, space technology, etc. These flows can be achieved by either time dependent movement of the boundary or unsteady temperature of the boundary. The unsteady temperature may be attributed to the free stream oscillations or oscillatory flux or temperature oscillations. The oscillatory convection problems are important from the technological point of view as the effect of surface temperature oscillations on skin friction and heat transfer from surface to the surrounding fluid has special interest in heat transfer engineering.

Flows which arise due to the interaction of the gravitational force and density differences caused by the simultaneous diffusion of thermal energy have many applications in geophysics and engineering. Such thermal and mass diffusion plays a dominant role in a number of technological and engineering systems. Obviously, the understanding of this transport process is desirable in order to effectively control the overall transport characteristics. The problem of combined buoyancy driven thermal and mass diffusion has been studied in parallel plate geometries by a few authors, notably, Lai[1], Chen et al.,[2], Mehta and Nandakumar[3] and Angirasa et al.,[4].

Adrian Postelnicu [5], Emmanuel Osalusi et al.,[6], Mohammed Abd-El-Aziz[7] have studied thermo-diffusion and diffusion thermo effects on combined heat and mass transfer through a porous medium under different conditions.

Theoretical study of free convection in a horizontal porous annulus, including possible three dimensional and transient effects. Similar studies for fluid filled annuli are available in the literature [8]. In view of this, several authors, notably Tunc et al [9],Oliveira et al.,[10]. Martin Ostoja [11], El – Hakein [12], and Bulent Yesilata [13] have studied the effect of viscous dissipation on convective flows past an infinite vertical plates and through vertical channels and ducts.

II. The Problem formulation

We consider the motion of viscous, incompressible, electrically conducting fluid through a porous medium in a vertical channel bounded by flat walls. The thermal buoyancy in the flow field is created by a traveling thermal wave imposed on the boundary wall at $y = L$ while the boundary at $y = -L$ is maintained at constant temperature T_1 . The walls are maintained at constant concentrations. The Boussinesq approximation is used so that the density variation will be considered only in the buoyancy force. The viscous and Darcy dissipations are taken into account to the transport of heat by conduction and convection in the energy equation. We take Soret effect into account in the diffusion equation. Also the kinematic viscosity ν , the thermal conductivity k are treated as constants. We choose a rectangular Cartesian system $O(x,y)$ with x -axis in the vertical direction and y -axis normal to the walls. The walls of the channel are at $y = \pm L$. The equations governing the unsteady flow and heat transfer are

Equation of linear momentum

$$\rho_e \left(\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} \right) = - \left(\frac{\partial p}{\partial x} \right) + \mu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) - \rho g - \left(\frac{\mu}{k} \right) u - \left(\sigma \mu_e^2 H_0^2 \right) u \quad (2.1)$$

$$\rho_e \left(\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} \right) = - \left(\frac{\partial p}{\partial y} \right) + \mu \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) - \left(\frac{\mu}{k} \right) v \quad (2.2)$$

Equation of continuity

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 \tag{2.3}$$

Equation of energy

$$\rho_e C_p \left(\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y}\right) = \lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) + Q + \mu \left[\left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2\right] + \left(\frac{\mu}{\lambda k} + \sigma \mu_e^2 H_0^2\right)(u^2 + v^2) \tag{2.4}$$

Equation of

Diffusion

$$\left(\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y}\right) = D_1 \left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2}\right) + k_{11} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) \tag{2.5}$$

Equation of state

$$\rho - \rho_e = -\beta \rho_e (T - T_e) - \beta^* \rho_e (C - C_e) \tag{2.6}$$

where ρ_e is the density of the fluid in the equilibrium state, T_e, C_e are the temperature and concentration in the equilibrium state, (u, v) are the velocity components along $O(x, y)$ directions, p is the pressure, T, C are the temperature and concentration in the flow region, ρ is the density of the fluid, μ is the constant coefficient of viscosity, C_p is the specific heat at constant pressure, λ is the coefficient of thermal conductivity, k is the permeability of the porous medium, D_1 is the molecular diffusivity, k_{11} is the cross diffusivity, β is the coefficient of thermal expansion, β^* is the volumetric coefficient of expansion with mass fraction and Q is the strength of the constant internal heat source.

In the equilibrium state

$$-\left(\frac{\partial p_e}{\partial x}\right) - \rho_e g = 0 \tag{2.7}$$

where $p = p_e + p_D$, p_D being the hydrodynamic pressure.

The flow is maintained by a constant volume flux for which a characteristic velocity is defined as

$$Q = \frac{1}{2L} \int_{-L}^L u \, dy \tag{2.8}$$

The boundary conditions for the velocity and temperature fields are

$$\begin{aligned} u = 0, v = 0, T = T_1, C = C_1 & \quad \text{on } y = -L \\ u = 0, v = 0, T = T_2 + \Delta T_e \sin(mx + nt), C = C_2 & \quad \text{on } y = L \end{aligned} \tag{2.9}$$

where $\Delta T_e = T_2 - T_1$ and $\sin(mx + nt)$ is the imposed traveling thermal wave.

In view of the continuity equation we define the stream function ψ as

$$u = -\psi_y, v = \psi_x \tag{2.10}$$

Eliminating pressure p from equations (2.1) & (2.2), the equations governing the flow in terms of ψ are

$$\begin{aligned} \left((\nabla^2 \psi)_t + \psi_x (\nabla^2 \psi)_y - \psi_y (\nabla^2 \psi)_x\right) = \nu \nabla^4 \psi - \beta g (T - T_0)_y \\ - \beta^* g (C - C_0)_y - \left(\frac{\nu}{k}\right) \nabla^2 \psi - \left(\frac{\sigma \mu_e^2 H_0^2}{\rho_0}\right) \left(\frac{\partial^2 \psi}{\partial y^2}\right) \end{aligned} \tag{2.11}$$

$$\begin{aligned} \rho_e C_p \left(\frac{\partial \theta}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial \theta}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial \theta}{\partial y}\right) = \lambda \nabla^2 \theta + Q + \mu \left[\left(\frac{\partial^2 \psi}{\partial y^2}\right)^2 + \left(\frac{\partial^2 \psi}{\partial x^2}\right)^2\right] \\ + \left(\frac{\mu}{k} + \sigma \mu_e^2 H_0^2\right) \left[\left(\frac{\partial \psi}{\partial x}\right)^2 + \left(\frac{\partial \psi}{\partial y}\right)^2\right] \end{aligned} \tag{2.12}$$

$$\left(\frac{\partial C}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial C}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial C}{\partial y} \right) = D_1 \nabla^2 C + \left(\frac{ScS_0}{N} \right) \nabla^2 \theta \quad (2.13)$$

Introducing the non-dimensional variables in (2.11) - (2.13) as

$$x' = mx, \quad y' = \frac{y}{L}, \quad t' = t \nu m^2, \quad \Psi' = \frac{\Psi}{\nu}, \quad \theta = \left(\frac{T - T_e}{\Delta T_e} \right), \quad C' = \left(\frac{C - C_1}{C_2 - C_1} \right) \quad (2.14)$$

(under the equilibrium state $\Delta T_e = T_e(L) - T_e(-L) = \frac{QL^2}{\lambda}$)

the governing equations in the non-dimensional form (after dropping the dashes) are

$$\delta R \left(\delta (\nabla_1^2 \psi)_t + \frac{\partial(\psi, \nabla_1^2 \psi)}{\partial(x, y)} \right) = \nabla_1^4 \psi + \left(\frac{G}{R} \right) (\theta_y + N C_y) - D^{-1} (\nabla_1^2 \psi) - M^2 \left(\frac{\partial^2 \psi}{\partial y^2} \right) \quad (2.15)$$

The energy equation in the non-dimensional form is

$$\begin{aligned} \delta P \left(\delta \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial \theta}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial \theta}{\partial y} \right) &= \nabla_1^2 \theta + \alpha + \left(\frac{PR^2 E_c}{G} \right) \left(\left(\frac{\partial^2 \psi}{\partial y^2} \right)^2 + \delta^2 \left(\frac{\partial^2 \psi}{\partial x^2} \right)^2 \right) \\ &+ (D^{-1} + M^2) \left(\delta^2 \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2 \right) \end{aligned} \quad (2.16)$$

The Diffusion equation is

$$\delta Sc \left(\delta \frac{\partial C}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial C}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial C}{\partial y} \right) = \nabla_1^2 C + \left(\frac{ScS_0}{N} \right) \nabla_1^2 \theta \quad (2.17)$$

where

$$R = \frac{UL}{\nu} \quad (\text{Reynolds number})$$

$$G = \frac{\beta g \Delta T_e L^3}{\nu^2} \quad (\text{Grashof number})$$

$$P = \frac{\mu c_p}{k_1} \quad (\text{Prandtl number}),$$

$$D^{-1} = \frac{L^2}{k} \quad (\text{Darcy parameter}),$$

$$E_c = \frac{\beta g L^3}{C_p} \quad (\text{Eckert number})$$

$$\delta = mL \quad (\text{Aspect ratio})$$

$$\gamma = \frac{n}{\nu m^2} \quad (\text{Non-dimensional thermal wave velocity})$$

$$Sc = \frac{\nu}{D_1} \quad (\text{Schmidt Number})$$

$$N = \frac{\beta^* \Delta C}{\beta \Delta T} \quad (\text{Buoyancy ratio})$$

$$So = \frac{\beta^* k_{11}}{\nu \beta} \quad (\text{Soret Parameter})$$

$$M^2 = \frac{\sigma \mu_e^2 H_o^2 L^2}{\nu^2} \quad (\text{Hartman number})$$

$$\nabla_1^2 = \delta^2 \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

The corresponding boundary conditions are

$$\psi(+1) - \psi(-1) = 1 \tag{2.18}$$

$$\frac{\partial \psi}{\partial x} = 0, \quad \frac{\partial \psi}{\partial y} = 0 \quad \text{at } y = \pm 1 \tag{2.19}$$

$$\theta(x, y) = 1, \quad C(x, y) = 0 \quad \text{on } y = -1 \tag{2.20}$$

$$\theta(x, y) = \sin(x + \gamma t), \quad C(x, y) = 1 \quad \text{on } y = 1 \tag{2.21}$$

$$\frac{\partial \theta}{\partial y} = 0, \quad \frac{\partial C}{\partial y} = 0 \quad \text{at } y = 0 \tag{2.22}$$

The value of ψ on the boundary assumes the constant volumetric flow consistent with the hypothesis(2.8). Also the wall temperature varies in the axial direction in accordance with the prescribed arbitrary function t .

III. Shear Stress, Nusselt Number And Sherwood Number

The Shear Stress on the channel walls is given by

$$\tau = \mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)_{y=\pm L} \tag{3.1}$$

Which in the non- dimensional form reduces to

$$\tau = \left(\frac{\tau}{\mu U} \right) = (\psi_{yy} - \delta^2 \psi_{xx}) \tag{3.2}$$

$$= [\psi_{00,yy} + Ec\psi_{01,yy} + \delta(\psi_{10,yy} + Ec\psi_{11,yy} + O(\delta^2))]_{y=\pm 1}$$

And the corresponding expressions are

$$(\tau)_{y=+1} = b_{90} + \delta b_{91} + O(\delta^2) \tag{3.3}$$

$$(\tau)_{y=-1} = b_{92} + \delta b_{93} + O(\delta^2) \tag{3.4}$$

The local rate of heat transfer coefficient (Nusselt number Nu) on the walls has been calculated using the formula

$$Nu = \frac{1}{\theta_m - \theta_w} \left(\frac{\partial \theta}{\partial y} \right)_{y=\pm 1} \tag{3.5}$$

and the corresponding expressions are

$$(Nu)_{y=+1} = \frac{(b_{51} + \delta b_{52})}{(b_{44} - \sin(D_1) + \delta b_{45})} \tag{3.6}$$

$$(Nu)_{y=-1} = \frac{(b_{53} + \delta b_{54})}{(b_{44} - 1 + \delta b_{45})} \tag{3.7}$$

The local rate of mass transfer coefficient (Sherwood number Sh) on the walls has been calculated using the formula

$$Sh = \frac{1}{C_m - C_w} \left(\frac{\partial C}{\partial y} \right)_{y=\pm 1} \tag{3.8}$$

and the corresponding expressions are

$$(Sh)_{y=+1} = \frac{(b_{65} + \delta b_{63})}{(b_{58} - 1 + \delta b_{57})} \tag{3.9} \quad (Sh)_{y=-1} = \frac{(-b_{65} + \delta b_{63})}{(b_{58} + \delta b_{57})}$$

(3.10) where b_4, \dots, b_{90} are constants

IV. Discussion of the Numerical results

The aim of the analysis is to discuss the flow, heat and mass transfer of a viscous electrically conducting fluid through a porous medium in a vertical channel bounded by flat walls on which a travelling thermal wave is imposed. In this analysis, the viscous Darcy dissipation, Joule heating and Soret effect are taken into account. For computational purpose, we

take $P = 0.71$ and $\delta = 0.01$. It is observed that the temperature variation on the boundary, dissipative and Soret effects contribute substantially to the flow field. This contribution may be represented as perturbations over the mixed convective flow generated. These perturbations not only depend on the wall temperature, M , Ec and So but also on the nature of the mixed convective flow. In general, we note that the creation of the reversal flow in the flow field depends on whether the free convection effects dominates over the forced flow or vice versa. If the free convection effects are sufficiently large as to create reversal flow, the variation in the wall temperature, M , Ec and So affects the flow remarkably.

The variation of u with Soret parameter So shows that the reversal flow which appears in the vicinity of the left boundary disappears for higher $So > 0$ and $So < 0$. Also, $|u|$ depreciates with increase in $So > 0$ and an increase in $|So| < 0$, enhances $|u|$ in the left region and depreciates in the right region (Fig.1)

Fig.2 shows the an increase in $|So| > 0$ depreciates v in the entire flow region while in $|So| < 0$ enhances v in the left region and depreciates in the right region

An increase in $So > 0$ depreciates Rt in the flow region and an increase in $|So| < 0$ enhances Rt in the left region and reduces it in the right region (Fig. 3).

The non-dimensional temperature θ is shown in Fig.4 An increase in Sc or $So > 0$ enhances θ , while an increase in $|So| < 0$ depreciates the actual temperature .

The behaviour of C with Soret parameter So shows that an increase in $So > 0$ enhances the actual concentration and depreciates with $|So| < 0$ (Fig.5).

The shear stress on the boundary walls have been evaluated numerically for different G , Sc , and So , are shown in (Tables 1- 6) . Lesser the molecular diffusivity, lesser τ at $y = 1$ and larger τ at $y = -1$. An increase in $So > 0$ enhances τ in the heating case and depreciates it in the cooling case at $y = 1$ while enhances τ in both the heating and cooling cases with increase in $|So| (< 0)$. At $y = -1$, the stress enhances with $So > 0$ and depreciates with $|So| (< 0)$ for all $G (>, < 0)$ (Tables.1 and 2)

The average Nusselt number Nu which measures the rate of heat transfer has been exhibited in Tables. 3 and 4. The variation of Nu with the Soret parameter So reveals that $|Nu|$ at $y = 1$ enhances with increase in $|So| (> 0)$ and depreciates with $|So| (< 0)$ while at $y = -1$, it enhances with increase in $|So| (> < 0)$.

The Sherwood number Sh which measures the rate of mass transfer is shown in Tables.5 and 6 for different parametric values. The variation of Sh with Sc shows that lesser the molecular diffusivity, higher $|Sh|$ at $y = 1$ and lesser $|Sh|$ at $y = -1$ and lesser $|Sh|$ at $y = -1$. An increase in $|So| (> 0)$ depreciates $|Sh|$ at both the walls while an increase in $|So| (< 0)$ increases for $|G| = 10^3$ and depreciates for $|G| \geq 3 \times 10^3$ (Tables.5 and 6).

V. References

- [1] F. C. Lai, Int. commn. Heat Mass transfer, Coupled heat and Mass transfer by natural convection from a horizontal line source in saturated porous medium, 17 (1990) 489-499
- [2] T.S.Chen, C.F.Yuh, and A. Moutsoglou, Int. Journal of Heat Mass Transfer, Combined Heat and Mass transfer in mixed convection along vertical and inclined plates, 23 (1980) 527-537.
- [3] K.N.Mehta, and K.Nandakumar, Int. J. Heat Mass transfer, Natural convection with combined heat and mass transfer buoyancy effects in non-homogeneous porous medium, 30 (1987) 2651-2656.
- [4] D. Angirasa, G.P. Peterson, I.Pop, Int. Journal of Heat Mass Transfer, Combined heat and mass transfer by natural convection with opposing buoyancy effects in a fluid saturated porous medium, 40 (1997) 2755-2773.
- [5] Adrian Postelnicu , Int. Journal of Heat Mass Transfer, Influence of a magnetic field on heat and mass transfer by natural convection from vertical surfaces in porous media considering Soret and Dufour effects,47 (2004),1467-1472.
- [6] Emmanuel Osalusi, Jonathan Side, Robert Harris , Int. commn. Heat Mass transfer, Thermal diffusion and diffusion-thermo effects on combined heat and mass transfer of a steady MHD convective and slip flow due to a rotating disk with viscous dissipation and ohmic heating,35 (2008) 908-915.
- [7] Abd – El – Aziz Mohammed, Physics Letters A, Thermal diffusion and diffusion thermo effects on combined heat and mass transfer by hydromagnetic three dimensional free convection over a permeable stretching surface with radiation, 372 (2008) 263 – 272.
- [8] F.C.Lai, Int. commn. Heat Mass transfer, Coupled Heat and Mass Transfer by mixed convective from a vertical plate in a saturated porous medium, 18 (1991) 93 – 106.
- [9] G.Tunc,Y.Bayazitoglu, Int. J. Heat Mass transfer Heat transfer in microtubes with viscous dissipation, 44 (2001) 2395-2403.
- [10] P.J.Olive,P.M.Coelho,F.T.Pinho, J.Non-Newtonian fluid mech., The Graetz problem with viscous dissipation for FENE-P fluids, 12 (2004) 69-72.
- [11] Ostoja Martin – A.Starzewski, Int.Journal of Engineering and science, Derivation of the Maxwell-Cattaneo equation from the free energy and dissipation, .47(2009) 807-810.
- [12] M.A.El-Hakim, Int. commn. Heat Mass transfer, Viscous dissipation effects on MHD free convection flow over a nonisothermal surface in micropolar fluid, 27 (2000) 581-590.

[13] Bulent Yesilatha, Int. commn. Heat Mass transfer, Effect of viscous dissipation on polymeric flows between two rotating coaxial parallel discs, 29 (2002) 589-600.

Figures-Captions

Fig.1 u with S_0 , $Sc=1.3, N=1, M=2$

	I	II	III	IV
S_0	0.5	1.0	-0.5	-1.0

Fig.2 v with $S_0, Sc=1.3, N=1, M=2$

	I	II	III	IV
S_0	0.5	1.0	-0.5	-1.0

Fig.3 Rt with $S_0, Sc=1.3, N=1, M=2$

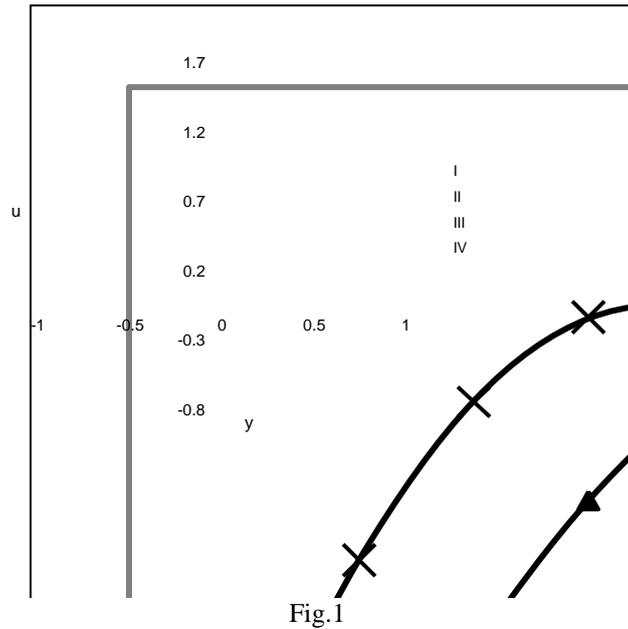
	I	II	III	IV
S_0	0.5	1.0	-0.5	-1.0

Fig 4 θ with Sc & $SoG=2 \times 10^3 m, D^{-1}=2 \times 10^3, M=2, N=1$

	I	II	III	IV	V	VI	VII
Sc	1.3	2.01	0.24	0.6	1.3	1.3	1.3
So	0.5	0.5	0.5	0.5	1.0	-0.5	-1

Fig.5 C with So

	I	II	III	IV
So	0.5	1.0	-0.5	-1.0



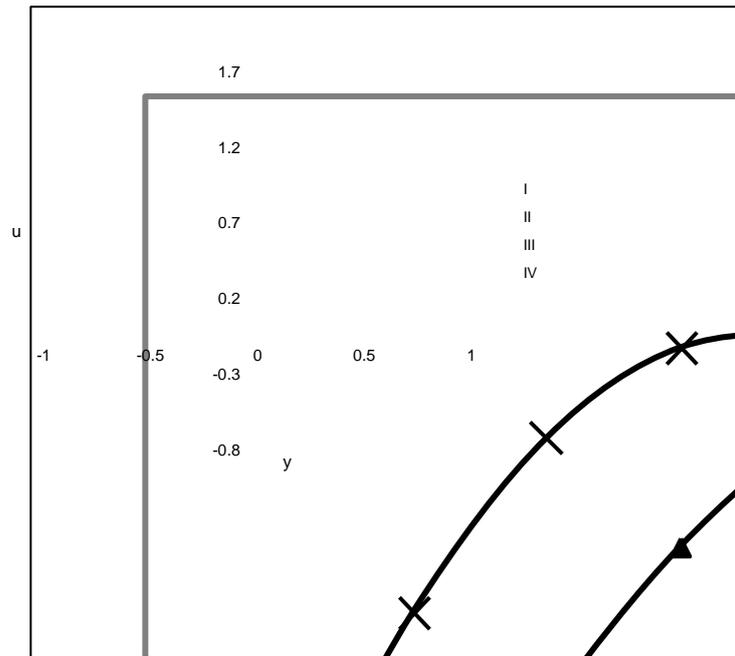


Fig.2

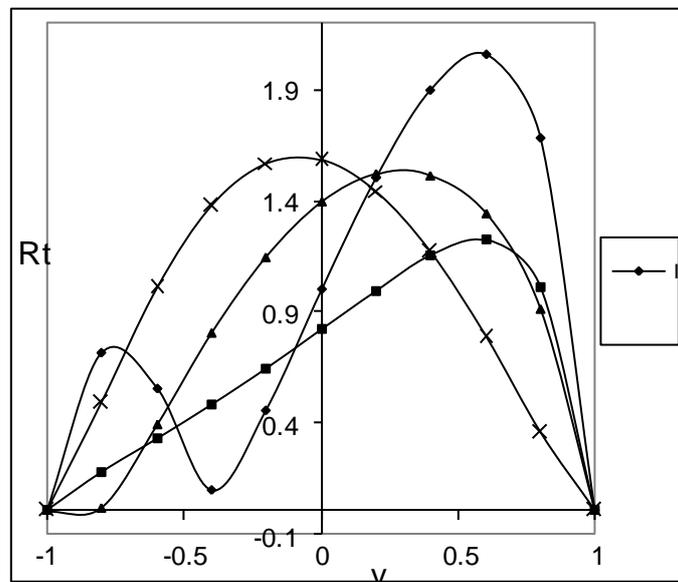


Fig.3

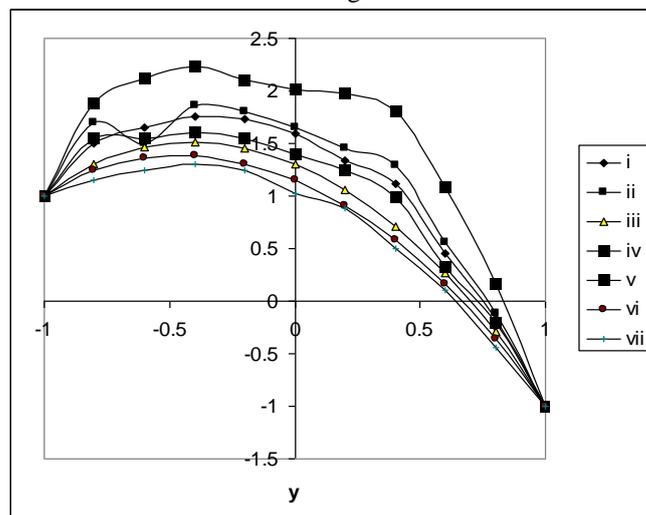


Fig.4

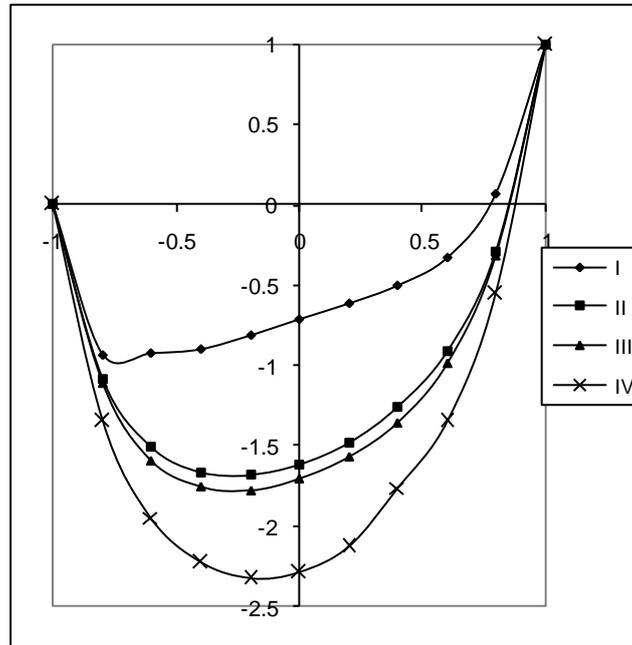


Fig.5

Table.1 Shear Stress (τ) at $y=1, P=0.71, x + \gamma t = \frac{\pi}{4}, D^{-1}=10^3, N=1, M=2$

G/ τ	I	II	III	IV	V	VI	VII
10^3	-9.6761	15.5287	-42.709	-27.539	-14.205	-151.32	-299.69
3×10^3	-10.132	11.7261	-265.16	-143.74	-104.85	-167.11	-312.17
-10^3	-11.038	19.0067	-49.949	-32.934	-6.8686	-162.09	-207.23
-3×10^3	-119.93	23.7171	-232.19	-362.41	-7.7614	-189.56	-286.78

Table.2 Shear Stress (τ) at $y=-1, P=0.71, x + \gamma t = \frac{\pi}{4}, D^{-1}=10^3, N=1, M=2$

G/ τ	I	II	III	IV	V	VI	VII
10^3	-1.0491	40.1131	0.3942	0.5227	-5.4294	-1.4265	-4.6217
3×10^3	-3.1121	34.1477	-4.5043	-3.2796	-9.6707	-18.304	-41.315
-10^3	5.2573	46.9346	6.7127	6.8696	11.5307	8.5845	1.6457
-3×10^3	9.3025	54.7654	-2.3945	2.2412	15.9049	-29.015	-59.501
	I	II	III	IV	V	VI	VII
Sc	1.3	2.01	0.24	0.6	1.3	1.3	1.3
S ₀	0.5	0.5	0.5	0.5	1.0	-0.5	-1.0

Table.3 Average Nusselt Number (Nu) at $y=1, P=0.71, x + \gamma t = \frac{\pi}{4}, N=1, M=2$

G/Nu	I	II	III	IV	V	VI	VII
10^3	-2.1681	-3.1421	-1.5071	-1.6211	-2.3371	-1.3321	-1.2846
3×10^3	-1.4014	-3.1502	-1.2211	-1.2351	-1.4641	-1.1055	-0.9654
-10^3	-2.1906	-3.1509	-1.5061	-1.6278	-2.3111	-1.3327	-1.2836
-3×10^3	-1.4036	-3.1479	0.12197	-1.2509	-1.4602	-1.1025	-0.9608

Table.4 Average Nusselt Number (Nu) at $y=-1, P=0.71, x + \gamma t = \frac{\pi}{4}, N=1, M=2$

G/Nu	I	II	III	IV	V	VI	VII
10^3	3.1504	2.9837	3.5057	3.4061	3.7321	3.1669	3.8022
3×10^3	3.5277	2.9858	3.7679	3.7241	3.9155	3.4642	4.1889
-10^3	3.1438	2.9839	3.5012	3.4006	3.7227	3.1131	3.8004
-3×10^3	3.5243	2.9863	3.7659	3.7224	3.9025	2.4685	4.1848
	I	II	III	IV	V	VI	VII
Sc	1.3	2.01	0.24	0.6	1.3	1.3	1.3
S ₀	0.5	0.5	0.5	0.5	1.0	-0.5	-1.0

Table.5 Sherwood Number(Sh) at $y=1$ $P=0.71$, $x + \gamma t = \frac{\pi}{4}$, $N=1, M=2$

G/Sh	I	II	III	IV	V	VI	VII
10^3	1.0619	1.5431	0.6162	0.5469	1.2133	2.2015	3.7554
3×10^3	0.0853	1.5356	0.5545	0.2526	-0.0693	1.8437	2.5289
-10^3	1.0805	1.5428	0.6179	0.5538	1.2019	2.1436	3.4643
-3×10^3	0.0854	1.5347	0.5548	0.2477	-0.0756	1.7516	2.2699

Table.6 Sherwood Number(Sh) at $y = -1$ $P=0.71$, $x + \gamma t = \frac{\pi}{4}$, $N=1, M=2$

G/Sh	I	II	III	IV	V	VI	VII
10^3	-1.3532	-1.5193	-6.7056	-1.9113	-1.3316	12.2447	24.1939
3×10^3	-0.4354	-1.5122	29.3535	9.8665	0.1615	-9.7093	-5.9196
-10^3	-1.3622	-1.5191	-6.7909	-1.9218	-1.3198	4.5883	67.6452
-3×10^3	-0.4562	-1.5119	19.4073	5.5267	0.1745	-7.0936	-4.5248
	I	II	III	IV	V	VI	VII
Sc	1.3	2.01	0.24	0.6	1.3	1.3	1.3
So	0.5	0.5	0.5	0.5	1.0	-0.5	-1.0

Through Transmission Laser Welding of Polymers using 1064 nm CW diode laser

Asst. Lect. Rana M. Taha

Laser & Optoelectronics Eng. Dep., College of Engineering, Al-Nahrain University, Baghdad, Iraq

Abstract: Through transmission laser welding provide high-speed, noncontact techniques for joining thermoplastics. A CNC machine provided by 1064 nm CW diode laser was used. Two sets of PMMA (Polymethylmethacrylate) specimens (Red with Black and Blue with Black) were welded. The mechanical and microstructure properties were measured. The shear strength was calculated for a clamping pressures of 3,5,7,9,&10 Bar with a welding speeds of 0.606 and 0.667 mm/sec where the second set of specimens (Blue with Black.) have a high shear strength than the first set under the same clamping pressure and welding speed conditions. The hole depth and width were calculated for the two sets and from this we observed that the increasing in the clamping pressure and welding speed affect the melting pool dimensions. The best results for the through transmission laser welding were achieved with a pressure range from 5 to 7 Bar and welding speed of 0.667 mm/sec.

Keywords: laser welding, PMMA, shear strength, clamping pressure.

I. Introduction

Thermoplastics are a group of polymers that can be classified according to their structure into two classes of amorphous and semi-crystalline polymers. They have applications in the automotive, aerospace, medical and electronics industries. Because of their excellent strength - to - weight ratio, they are chosen successfully for many applications where metal and ceramics have traditionally been used [1].

The using of laser in processing polymers has increased significantly. Infrared and excimer laser have been used for welding, cutting and modification of polymers as well as cutting of polymer-based resins. More recently, however, the unique properties of laser beams have been applied with spectacular success to the development of laser welding. The ability of lasers to generate the exact amount of energy and deliver it at speed to a well-defined location can be used for welding [2].

There are many different methods of joining plastics, with the selected process being dependent upon such variables as type of plastic, assembly requirements, and application area. These processes include welding, adhesive bonding, and mechanical fastening. Transmission laser welding involves localized heating at the interface of two pieces of plastic to be joined to produce strong, hermetically sealed welds with minimal thermal and mechanical stress, no particulates and very little flash. Cycle times can be as short as a second, and relatively light clamping pressure is required— just enough to keep the parts stationary and ensure there is no gap. Transmission

Laser welding can be used for rigid or flexible materials and small or large parts [3].

II. Through transmission laser welding:

It relies on the fact that many polymers, in their natural state, are not strong absorbers of NIR radiation. The window of NIR wavelengths that can be used, beyond the visible spectrum, from 800nm to approximately 2000nm, is illustrated in Fig.1.

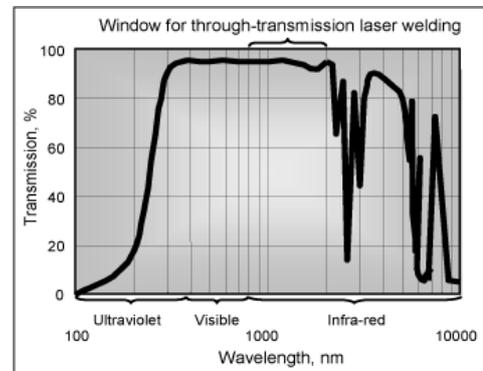


Fig.1. Absorption spectrum of a typical polymer[4]

The process differs from most other plastics joining processes, and from the use of lasers for joining metals, in that the energy from the laser is concentrated at the joint rather than melting material from the surface inwards. The process is illustrated in Fig.2. The energy of the laser must be concentrated at the joint in order to melt the polymer to form a weld. The laser passes through the upper workpiece to the joint. This melts a thin film of plastic on either side of the joint, resulting in a permanent weld when the joint has cooled [4].

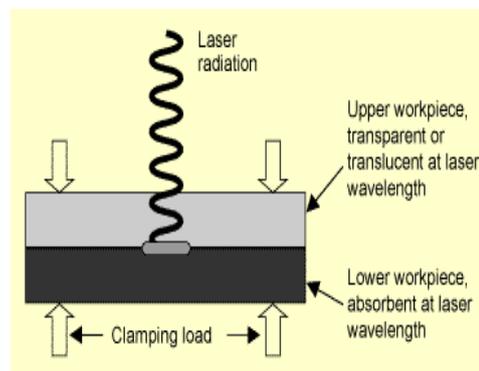


Fig.2. Through-transmission laser welding - cross-section through the welding process [4]

An important factor when using laser transmission welding is intimate contact between the surfaces to be joined. This ensures sufficient melt flow to produce a strong weld. To obtain intimate contact of the substrates at the weld interface, a certain amount of clamping pressure is required for most joints (the notable exception being an interference fit, as clamping may not be necessary). The amount of pressure depends upon the materials being joined, the specific joint design chosen, the quality of the surface conditions at the weld interface, and the final strength requirements. Surfaces typically have irregularities (bumps, valleys, etc.) which prevent close contact. As the surfaces heat and expand, the clamping pressure helps to flatten the surface, which removes entrapped air to impart even contact along the interface. This facilitates diffusion of the polymers necessary to create a strong weld [5].

III. Tensile test

The tensile strength of a material is the maximum amount of tensile stress it can be subjected to before failure. The definition of failure can vary according to the material type and design methodology. There are three typical definitions of tensile strength.

- Yield strength: the stress at which material strain changes from elastic deformation to plastic deformation causing it to deform permanently.
- Ultimate strength: the maximum stress a material can withstand.
- Breaking strength: the stress coordinate on the stress-strain curve at the point of rupture.

A lap shear joint was used for the polymers samples to examine the strength of the welded area, figure 3 shows an example of lap shear joint.



Figure 3: Schematic illustration (side view) of single lap joint before and after application of tensile stress [6].

IV. Experimental setup

Laser CNC machine:

A three dimensional Diode laser CNC machine was used and adapted to satisfy the welding process, as it was shown in fig. no.4. The diode laser parameters are 5W maximum power, CW mode at 1064 nm wavelength.

Clamping Pressure Device:

It was added to the laser CNC machine to adopt the machine for the laser polymer welding process.

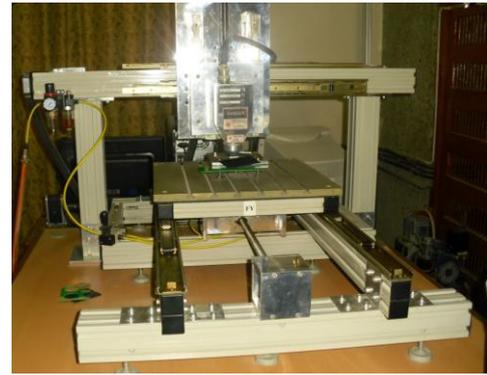


Fig. no.4: Laser CNC machine

Materials:

Three colours of 2.5 mm thickness PMMA (Polymethylmethacrylate) specimens were used, as shown in fig. 5, 6 & 7.



Fig. no.5: Red specimen



Fig. no.6: Blue specimen



Fig. no.7: Blue specimen

V. Absorption Spectrum:

One of the most important factor that must be considered in through transmission laser welding was the absorption of the used materials, fig. no. 8 , 9 & 10, shows the absorption spectrums of the three specimens.

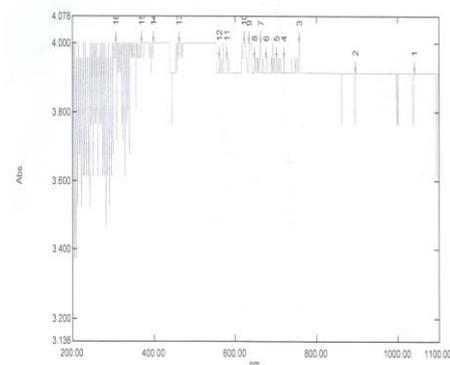


Fig. No.8: Absorbance of Black specimen

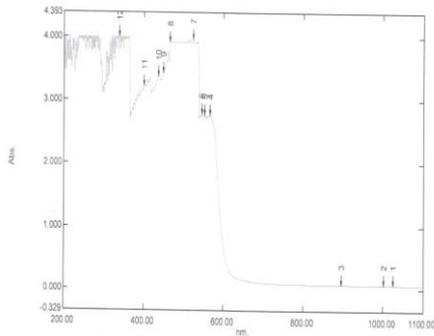


Fig. No.9: Absorbance of Red specimen

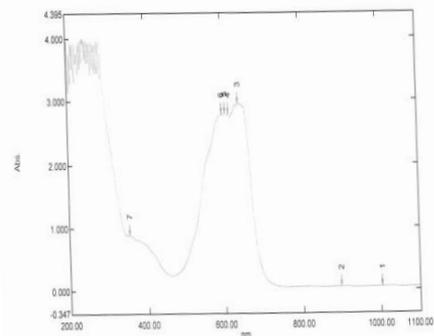


Fig. No.10 : Absorbance of Blue specimen

From fig. no.8, we observed that the black specimen has a high absorbance at 1064 nm wavelength which reaches to 3.9, while for the red and blue specimens, the absorbance reaches to 0.065 and 0.048 respectively.

VI. The Procedure

A two sets of PMMA specimens (Red with Black and Blue with Black) were welded, , fig. no. 11 & 12 shows the two welded sets.



Fig. no.11: 1st set of welded specimens

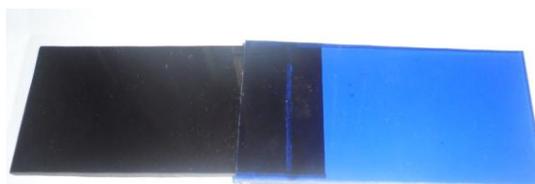


Fig. no.12: 2nd set of welded specimens

VII. Shear test:

Strength testing was an important aspect of a weldability study in laser welding, and among all tests, static tensile-shear testing was the most common laboratory

test used in the determination of weld strength because of its simplicity in which the average forces were calculated to each couples of specimens then the shear stress were calculated using the following equation:

$$\tau_{sh} = \frac{P}{A} \text{ (N/mm}^2\text{)}$$

Where; P=Failure Force (N)

A= Spot Area = πab (mm²)

a, b= Spot Radii (where the spot have an ellipse shape)

Fig. no. 13 and 14 shows the shear strength versus the clamping pressure for two sets of specimens.

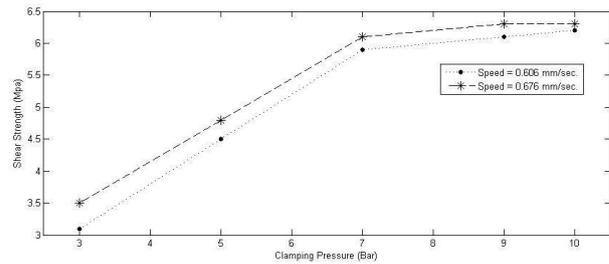


Fig. No.13: Shear strength (Mpa) versus Clamping pressure (Bar) of the 1st set

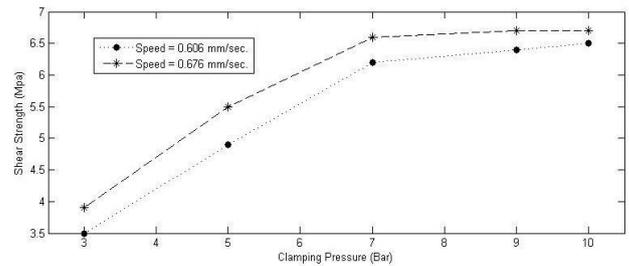


Fig. No.14: Shear strength (Mpa) versus Clamping pressure (Bar) of the 2nd set

From fig. no.13 & 14, we noticed that the shear strength was increased with an increasing of the clamping pressure and from that figures we noticed that the maximum increasing in the shear strength occurs in clamping pressure between 5 to 7 bar also we can noticed that the values of the shear strength of the second set (blue with black) have a larger values than that of the first set (Red with Black) and this is because the red specimen have a larger value of absorbance than that of blue specimen.

Fig. No. 15 & 16, shows the upper depth of the two sets of specimens.

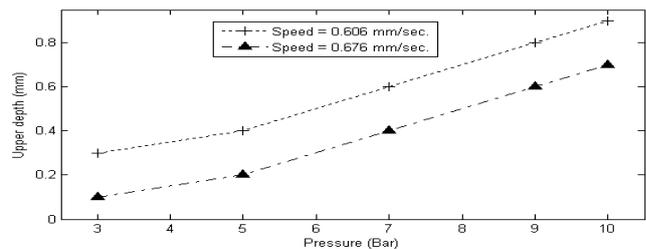


Fig. No.15: Upper depth (mm) versus Clamping pressure (Bar) of 1st set

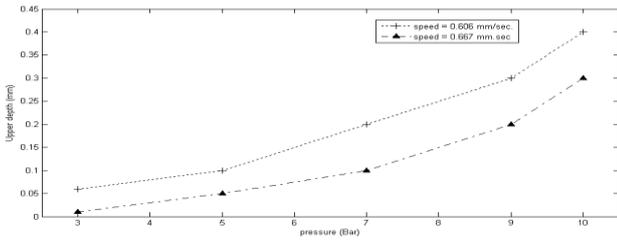


Fig. No.16: Upper depth (mm) versus Clamping pressure (Bar) of 2nd set

From fig. no. 15 & 16, we noticed that the upper depth was increased with an increasing in the clamping pressure also it was affected by an increasing in welding speed where it was inversely proportional with it, also the second set have a relatively lower values for upper depth than that of the first set.

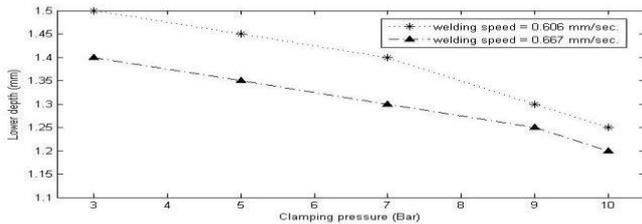


Fig. No.17: Lower depth (mm) versus Clamping pressure (Bar) of 1st set

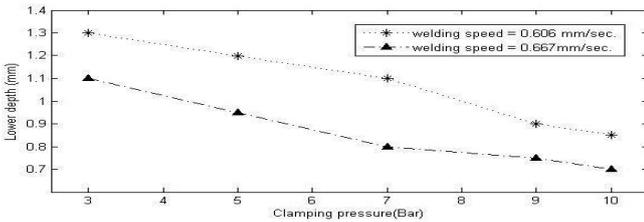


Fig. No.18: Lower depth (mm) versus Clamping pressure (Bar) of 2nd set

From fig. no. 17 & 18, we noticed that the lower depth was inversely proportional with an increasing in the clamping pressure and welding speed, and as in the upper depth the lower depth also have relatively a lower values than that of the first set of specimens.

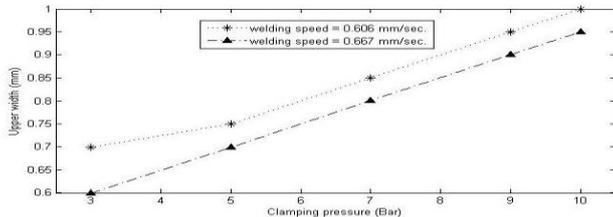


Fig. No.19: Upper Width (mm) versus Clamping pressure (Bar) of 1st set

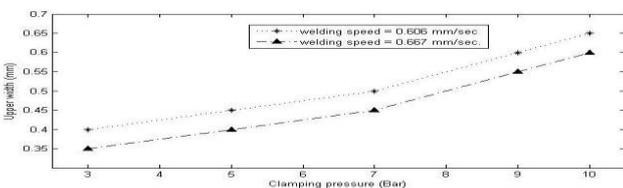


Fig. No.20 : Upper Width (mm) versus Clamping pressure (Bar) of 2nd set

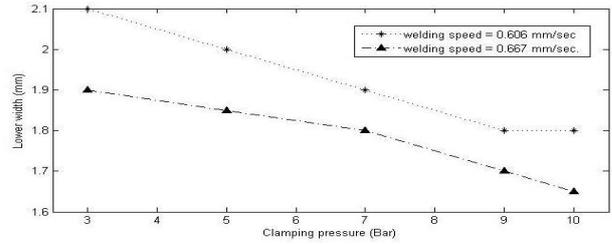


Fig. No.21 : Lower Width (mm) versus Clamping pressure (Bar) of 1st set

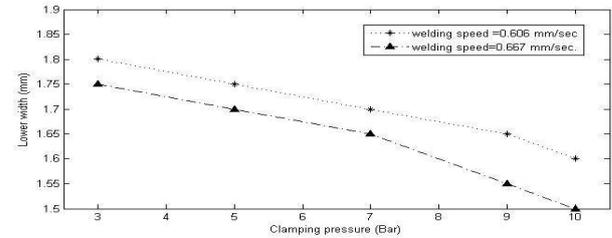


Fig. No.22 : Lower Width (mm) versus Clamping pressure (Bar) of 2nd set

From fig. no. 19, 20, 21 & 22, we noticed that the upper and lower width were also affected by increasing in clamping pressure and welding speed and they were taking the same platform of upper and lower depth.

VIII. Conclusion:

1- From fig. no. 13 & 14, we can conclude that:

- * The clamping pressure must have a limit in increasing where when it reached to a certain value, the increasing in the shear strength becomes slightly.
- * The second set of specimens (Blue with Black) have a better welding quality that the first one (Red with Black) .
- * The best results for shear strength of two sets were achieved at clamping pressure between 5 to 7 bar and welding speed of 0.667 mm/sec.

2- From figures no.15 through 22, we can conclude the clamping pressure and welding speed play a very important role to enhance the mechanical properties of the welded zone.

References

- [1]. M. R. Nakhaei, N. B. Mostafa Arab, F. Kordetani, "Modeling of weld lap-shear strength for laser transmission welding of thermoplastic using artificial neural network", *Advanced material research*, vol.445, p 454-459, 2012.
- [2]. Gareth C McGrath, Nicole Woosman and L P Frieder III, "Latest development in the joining of polymers for use in medical applications", *Medical polymers 2003*, Dublin, Ireland, 2-3 April 2003.
- [3]. Welding technology institute of Australia, "Transmission Laser Welding of Plastics", www.wtia.com.au.
- [4]. Marcus Warwick and Marcus Gordon, "Application studies using through-transmission laser welding of polymers", *Joining Plastics 2006*, London, National Physical Laboratory (NPL), 25-26 April 2006.
- [5]. Nicole M. Woosman and Michelle M. Burrell, A Study of the Effect of Weld Parameters on ClearweldedTM Thermoplastics, Gentex Corporation, Carbondale, Pennsylvania, USA.
- [6]. Single Lap-shear testing, Retrieved on 30/2/2007 from www.sintef.com.

Compositional Basis of Biological Design

Andrew Kuznetsov

University of Freiburg, ATG:Biosynthetics, Germany

Abstract: *Asymmetric sophisticated interactions between parts of a system that is embedded in the environment which changes from time to time leads to modularisation and to growing complexity because the environment has also evolved as a result of the activity of its inhabitants. The author has analysed research on modularity in biology and his own results in this field. Examples of a complex system and programmable self-assembling are given. Recognition and exploitation of modularity in artificial and natural systems are demonstrated. A particular role of horizontal gene transfer as a novel hypothesis on biological modularity is given special attention to. A formal model of compositional evolution under supervision of the environment that changes over time is considered. Some examples of the bottom-up design for artificial genetic networks of increasing complexity are obtainable. Different kinds of behaviour for virtual creatures – sperm cells and ova – are shown. Presented case studies may be useful for the development of a common theory of biological design.*

Keywords: *non-linear dynamics, biological complexity, modules, gene transfer, emergences*

I. Introduction

Modularity is an old concept in biological science, which is based on the ancient belief that modularity was a real and universal property of Nature. In the 18th century, comparative anatomists such as Georges Cuvier and Geoffroy Saint-Hilaire identified structural modules representing parts of organisms. Joseph Needham in the 1930's proposed that development can be decomposed into separate elements [1]. In modern times, Walter Fontana, Günter Wagner, Uri Alon, and many others have contributed much to the field [2-6]. Here is a list of recently discovered rules:

- A constant environment that does not change over time leads to non-modular structures
- In contrast, the modular structure can spontaneously emerge if environment changes over time
- Variability in the natural habitat of an organism promotes modularity
- Modularity can also dramatically speed up evolution
- Adaptation of bacteria to new or changing environments is often associated with the uptake of foreign genes through horizontal gene transfer (HGT).

My own research suggests that HGT is the important force, which contributes significantly to modularity [7]. The living systems contain detectable modules such as gene clusters, protein domains, organelles, cells, and tissues and organs with specific biological functions. Just as Richard Watson has written: “the existence of modularity in Nature is now becoming testable” [8], the modularity was revealed in complex biological networks and regulation systems [9-13]. In particular, correlation in high-throughput data and phylogenetic profiles were used for detecting functional modules [14-16].

The main goals of this paper are (1) to assign a role of modularization as the minimalist economic principle for the adaptation to variable environment, (2) to stress the significance of the interaction between modules, and (3) to underline the role of horizontal gene transfer (HGT) in adaptation and evolution because HGT was a major effort in my “evolution by communication” study [17]. We conjecture from previous results that modularity will help to answer the following questions: How can evolution lead to modular systems? Why does modularity exist in biology? Why do complex systems fail or die?

The article is divided into five major parts. The first part is a brief introduction to the field of complexity; the fundamental problem of science with an example of complex system behavior, as well as demonstrations of self-assembly. The second part of the manuscript is about recognition of modules in artificial reality and natural systems. In addition, a case study will be presented of how an inspiration from natural modularity can help resolve real technical challenges. The third part is dedicated to answering a crucial question concerning the origin of modularity based on the analysis of distributed bacterial sulfur metabolism. This section is about the nature of mutations and describes a generalization of copy, cut, and paste mechanisms. The fourth part of the paper looks at the bright possibilities of modular design. It will bring the modularity of genetic networks in *pi*-calculus with examples of elements of networks and genetic motifs. Finally, the ideas of scalable design, emergent behavior, and compositional evolution of biological systems will be considered. Readers will be invited to a speculation about the role of horizontal gene transfer in the biological complexity. The main point presented in this document is the significant role of the transfer and interaction of modules in adaptation, which is still in the shadow of the general evolution paradigm. An example of the program code is included in the Appendix.

II. Complexity

Let me begin with a bit of philosophy. A fundamental problem in science is why matter grows in complexity. As Herbert Simon wrote in the introduction to the book ‘*Modularity: understanding the development and evolution of natural complex systems*’ [18], “Complexity arises then ... components interact with each other in ways ... more than uniform, frequent elastic collisions. Interactions among components can lead to all kinds of nonlinear behavior.” The phase space of complex systems usually exhibits irregular surfaces of local minima and maxima, or even demonstrates bifurcations and chaotic behavior. According to Simon, if we take the phrase “survival of the fittest” literally, then the theory of evolution has

validity only in a world where maxima are attainable and the paths toward them are discoverable, which is seldom the case in the real world.

As an example, I wish to present some amazing data obtained when I was a student, the divergence of astrocytes for GFA content depending on malignation [19, 20]. GFA, which means Glial Fibrillary Acidic Protein, is a neurospecific protein, which is a marker of nervous cells such as astrocytes. It is the biomarker of the neoplasia of the Central Nervous System (CNS). We investigated a distribution of GFA protein for different cancer cells in CNS. Because there was a divergence of nervous cells for GFA protein content depending on malignation, which was not easy to diagnose, the linear and nonlinear regression functions were applied gradually. Distributions of soluble and insoluble forms of GFA protein in the normal and malignant nervous cells were finally well described by the canonical cusp catastrophes, i.e. Whitney cusp surfaces (Figure 1).

Unfortunately, ordinary mathematical methods no longer lead to solutions in closed form and moreover, the complexity can carry us beyond the simulation capacities. Examples of 3D and 2D fractal structures in Figure 2 are results of the diffusion limited aggregations (DLA) generated within a computer. Figure 3 presents screenshots of the assembling by adhesion rules. The exercise was to form a chessboard pattern in the presence of DLA. The seed of crystallization is in the bottom-left corner. Figure 3a demonstrates a random initial position of tiles and Figures 3b-d present the final solutions for different parameters of diffusion [21]. One can see that it is impossible to create the chessboard pattern without gaps, but Nature does it better. The great problem is to find those tricky rules of interactions between soft living particles, for instance repair rules, examples of robustness, and so on.

With respect to Herbert Simon's point of view on complexity, I would like to recall his definition of modularity in the terms of interactions. He wrote: "... the frequencies of interaction among elements in any particular subsystem of a system are an order of magnitude or two greater than the frequencies of interaction between the subsystems. We call this ... nearly decomposable (ND) system." [22]. In other words, ND systems are made up of separate parts where there is far more relations within each part than between different parts. ND is not the same as modularity but it gives a clue about an essential property of any modular system; this model is very general. Simon continued later: "A system may be characterized as modular to the extent that each of its components operates primarily according to its own, intrinsically determined principles. Modules within a system or process are tightly integrated but relatively independent."

III. Recognition of modules

Examples of modularity are quite natural in today's software design, everyday engineering practice like electronics, optics, and even in DNA-nanotechnology such as DNA-origami building blocks, modular DNA folding, and DNA-protein interactions in the case of binding complex GCN4 bZIP with DNA [23]. A modular design facilitates development and maintenance of complex technical devices. In natural science, modularity is represented by a search for fundamental units and basic elements, for instance, elementary particles, chemical elements, molecules and compounds, gene clusters, metabolic pathways, etc. The search for modularity is the identification of sets of base elements and construction rules to recognize simplicity in complex systems.

Figure 4 demonstrates an example in nanotechnology, the result of computer added design of endohedral metallofullerenes on the basis of quantum mechanical calculations within the density function theory (DFT). Cobalt-clusters and carbon-fullerenes were considered as independent modules which can be in different positions and could interact in different ways. We were able to use this approach to plan experiments rationally and to find an empirical rule for the interaction between Co- and C-atoms that describes the magnetic moment of the complex as a function of number and length of chemical bonds: the magnetic moment M per Me-atom of given complex is proportional to the average Me-C bond length L divided by the total number N of Me-C bonds in the complex [24, 25].

2.1. What is a module? As was shown above, the idea of modularity is intuitive but is hard to formulate. To give an example; Uwe Strähle and Patrick Blader wrote: "... we define a module as an assembly of biological structures that fulfill a function in an integrated and context insensitive manner. Function as defined here is not merely the interaction of molecules but an interaction that yields a biological output which is characteristic of the module. Furthermore, the application of the module is flexible. To be recognized as a module, it has to be used either in different processes in the same organism or in different organisms, exploiting its invariant functional properties in the same or different processes. A module is therefore characterized by its reiterated use." [26]. It looks rather difficult. I am sure some readers have recognized this problem. In view of Uri Alon having defined modularity in a more laconic way; a "property of a system which can be separated into nearly independent sub-systems" [27]. In other words, modularity is defined through a process that starts by recognizing patterns, shapes, or events that are repeated at some scale of observation. Modularity is a hallmark of biological organization and an important source of evolutionary novelty. Modularity is a sign of the universal principle of economy in Nature.

Biological systems present both genotypic and phenotypic modularity. Most biological functions are carried out by particular groups of genes and proteins so that one can split the structure into functional modules. For example, proteins work in groups, such as complexes and pathways. I like this simple definition: *a module is a set of genes that act together to carry out a specific function*. The modularity of biological networks is puzzling and the recognition of modularity came as a surprise. In this situation: (1) find modules, relations between modules, the origin of modules, (2) understand the hierarchy of a modular system and a reason of the entanglement within modules and between modules because modularity is the basis of our ability to separate problems into smaller parts that can be studied independently to assign functions to genes, proteins, metabolic and signaling pathways. In my opinion, the answers to the following questions could have given a key to control an evolutionary process: (1) How does a system evolve and fail? (2) What is the limit of evolvability? In addition, I would

like to point out that *evolvability* is the ability to respond to a challenge by producing the correct variation.

2.2. Natural modularity (*dsr* and *sox* gene clusters). I investigated *dsr* and *sox* gene clusters coding enzymes for bacterial community sulfur metabolism (*dsr* means the dissimilatory sulfate reduction and *sox* means the sulfide oxidation). Studies on environmental DNA databases such as Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis – CAMERA for short, allowed me to confirm the major role of HGT in modularity. The correlation between the *dsr* and *sox* clusters for the experimental set of 41 stations around the world was $R = 0.86$, which demonstrates the complementarity of *dsr* and *sox* metabolic pathways in environmental populations. Genes tend to group in modules that facilitate the propagation of specific functions within a community of organisms. To confirm this observation, the AprA tree was produced using *A.pompejana* symbiont reference and NCBI BLAST pairwise alignments. However, it was impossible to map AprA tree to 16S rRNA phylogenetic tree because putative HGT affected the canonical phylogenetic tree. I revealed a hierarchical modularity in the bacterial sulfur metabolism. The example included a repressor of phase-1 flagellin within the large sulfite reductase 4Fe-4S domain [7].

IV. Origin of modularity

According to the Günter Wächtershäuser, William Martin, and Eugene Koonin scenario of the origin of life [28, 29], the sulfur metabolism was a very ancient invention of Nature. The data obtained from modern DNA-sequence databases allowed me have a look back 3.5-3.8 billion years, when unicellular life emerged on our planet. This logic helped me to merge the origin of life with the origin of modules. Modularity could be explained as an adjustment for evolvability. I concluded that modularity is an unavoidable design feature of organic life [7]. Evidences of HGT and modular nature of prokaryotic genomes are very good news for engineers who are trying to design complex metabolic pathways. Such artificial modular designs might include a kernel (housekeeping genes) and modules involved in a process of adaptation toward a particular environment. Travelling groups of genes could be easily embedded in almost any system.

3.1. Nature of mutations. Unfortunately, biological reality is much too complex to be captured by a linear mapping of genes to phenotypes. However, it is a reasonable mechanistic assumption that the genome produces adaptive variants. Such a genome is able to get access to regions of the phenotype that may be adaptive in two ways: (1) by the mechanisms of compositional evolution which combine interdependent genetic modules that have evolved previously in parallel, and (2) using small gradual changes like point mutations. In other words, it means (1) the change of topology of genetic networks, and (2) the change of parameters which are induced in DNA sequences.

An example from experiments on the CAMERA database is in Figure 5, the alignment of the protein reads from microorganisms living in the gutless worm *O.algarvensis* to the AprA polypeptide from bacterium *V.okutanii*. I was able to show different patterns of variability for AprA protein [7]. Numerous exchanges, deletions and insertions in AprA polypeptide were found as are marked in yellow.

3.2. Modularity as a set of construction rules, the cut and paste Argo-machine. An extreme specialization is somewhat risky for an organism in diverse conditions. A remodularization of the organism is needed in the changing environment. Consider an evolving system – an abstract machine and an environment that is continuously changing creates advice words for the machine to stimulate an adaptation of this device to its surroundings (Figure 6a). The input for this machine is special words which are generated by the environment. We call these input words ‘oracle’, ‘guide’, or ‘generative’. The output is phenotypes which fit the environment. The machine operates on strings, which can code organisms in hierarchical manner. This non-deterministic abstract machine searches according to oracle words in the design space proper to its environment by cutting, transposing, and pasting a set of tapes [30].

The schematic of *Argo-machine* (*AM*) with circular tapes is shown in Figure 6d. The *AM* consists of agents and each of these has a head and a tape and can be in different output states. In general, the tape is a nonempty string of symbols that may be linear or circular. The head scans the tape according to an input word w_i and cuts it at recognized sites. The agent arbitrarily pastes the tape. For each tape-configuration there is an appropriate output state of the agent that is checked by the environment. Special ‘accept’ and ‘reject’ states take immediate effect. An agent accepts if its output state corresponds to the environment state; an agent will reject if less than two matches to the input word exist on the tape. *AM* can accept if at least one agent accepts, reject if all agents reject or loop (Figure 6b). If the environment has changed, it delivers a transposition and a new word w_{i+1} . The transposition means to make a copy of the tape from the accepted agent to other ones and join it from head-to-tail (Figure 6d). *AM* continually looks for an agreement with the environment.

We could describe *AM* in another way. The system operates on inputs and memory, uploads the memory, and yields outputs. Changes in environment generate ‘oracle’ words which guide DNA shuffling and transpositions. A combinatorial power of this system is very high. In short, *AM* is a set of stochastic cut-paste agents, which act in parallel on their own tapes according to the instructions (input words), communicate with each other by transpositions of the tape, and interact with the environment to compare the output states. Based on the comparison it accepts or runs in a loop to fit to the environment.

A crucial point of *AM* is the *Argonaut algorithm*. Each agent, Argonaut in other words, performs the next actions on word w :

- (1) Scans the tape to be sure that it has at least two matches. If not, rejects it.
- (2) Cuts at the matching sites and arbitrarily paste the tape’s fragments.
- (3) Takes the output state according to the new tape.
- (4) Checks the output state with the state of the environment. If satisfied, accepts; otherwise loops.

The combinatorial power of the *Argonaut algorithm* is bigger than polynomial, exponential, and factorial functions (Figure 6c). Computation associated with *Argo-machine* is the shuffling of tapes from the initial set T_0 until an accepted state is reached – an adaptation. Usually, computation never ends, because the environment changes permanently; if this happens, the case, called a catastrophe, leads to a transposition, generates a super-transition from the accept-state to the set of new initial-states, and brings a new generative word. *A progression of adaptations and catastrophes is an evolution.*

V. Modularity of genetic networks in *pi*-calculus, a modular “table of elements”

The model in our case is an abstraction that summarizes the property of a modular structure. Some people believe that Nature uses a restricted number of models to establish a set of relations among modules. I used this *à la* Wolfram approach, introduced in 2002 by Stephen Wolfram in his book [31]. In contrast to Wolfram’s 1D-cellular automata, my research was based on *pi*-calculus to build genetic networks. The *pi*-calculus was invented in 1992 by Robin Milner with the aim of describing interactive systems [32]. As a next step, the stochastic *pi*-calculus was proposed by Corrado Priami in 1995 [33], while Andrew Phillips from Microsoft developed SPiM language and designed a stochastic simulator (Stochastic Pi-Machine, SPiM for short) that performed a Gillespie algorithm [34, 35]. I used SPiM language to write a code for the SPiM simulator (see Appendix).

The following primitives in SPiM calculus were used to make networks with different topologies: decay (degradation of a transcription factor), *null* gate (constitutive transcription), gene product (protein transcription factor), *neg* gate (negative regulation), *pos* gate (positive regulation). I built genetic circuits of increasing complexity from those five basic primitives. Furthermore, I investigated the circuits such as negative and positive genetic regulators, genetic switches, oscillators, impulse generators, feedback and feedforward loops, genetic memory elements, and bifan motifs.

My work in the ‘silicon laboratory’ represented the bottom-up scenario of compositional design in terms of SPiM calculus. Most models chosen arbitrarily worked well at standard values of parameters. The topology and complexity of networks played a significant role in their behavior. New networks emerged easily, sometimes because of duplications and transpositions of the earlier ones. These operations resembled natural genetic mechanisms such as vertical and horizontal gene transfer – essential operators of biological evolution. Some variations in parameters were similar to ‘point mutations’ leading to an optimization (adaptation) of the system to a desirable pattern of behavior [36].

4.1. Basic genetic gates. Figure 7 presents the basic genetic gates, such as negative and positive regulators without input, with regulated input, also the cases of autoregulation. For each plot, an abscissa indicates the time of simulation with an ordinate that is the number of protein molecules. Simulations were started in the absence of proteins by doing a constitutive transcription. The number of protein molecules initially increased along with the time of simulation and finally levelled off at the equilibrium between production and degradation. The constitutive expression and the output were higher for the negative regulator than for the positive one.

To see a response of genetic elements, an input allowed linear increases from 0 to 100 individual molecules then decreased linearly to 0. The input molecules were injected into the system at a certain times. As a result of the reaction, the negative gate behaved like an inverter whereas the positive gate increased the output signal almost 10 times. The results of negative and positive autoregulations are also shown in Figure 7.

4.2. Repressilator. A repressilator consists of three *neg* gates that mutually repress each other as shown on the picture (Figure 8). Simulation of the repressilator at nominal parameters resulted in an irregular duration of protein cycles. The decreased rate of gene unblocking $\eta_p=0.001$ and the increased protein binding $r=10.0$ allowed an improvement of the regularity of oscillations. Populations of proteins stabilized nearly 100 molecules in each cycle with the same duration of impulses. The program code for this experiment is given in the Appendix.

4.3. Bi-stability and memory. Previous experiences were used to design a genetic memory element. The closed chain from four *neg* elements demonstrated bi-stable characteristics. The system arbitrarily started from expression (*a,c*) or (*b,d*) proteins. Fortunately, the circuit established stable behavior at a standard range of parameters. After it dropped to an arbitrary state, the system survived for a long time (Figure S1a,b, insertions). However, the circuit was sensitive to external inputs. For example, when the system is in state *bd*, then a programmable input *a* can change its state to a new state *ac*, to be exact, the production of *b* and *d* proteins can be changed to *a* and *c* proteins by input *a* (Figure S1c). Moreover, if input *a* no longer exists, the system nevertheless stays at the state *ac* and does not turn back until the specific input *b* changes the system’s state again (Figure S1d).

4.4. Synchronous FFBL. I discovered that the coherent feedforward and feedback loops (FFBL) circuit can be in four different states corresponding to particular patterns of protein expression: I, red – low basal production of *b* and *c* proteins, II, green – intensive stable production of *b* and *c*, III, blue – spontaneous synchronous outputs of *a* and *c*, and IV, black – exhaustive expression of *a* and *c* proteins with gaps. These essential states were significantly influenced by stochastic fluctuations. Nevertheless, I did not observe any transition from one state to another when the parameters were fixed (Figure S2).

4.5. Asynchronous FFBL. An incoherent FFBL circuit showed even more sophisticated outputs. Sometimes the system demonstrated unstable dynamic behavior. I found this circuit in the following states: I, red – decoherent small amounts of *b*

and large amounts of c proteins, II, green – low level of b protein and intermediate level of c protein, III, blue – flip-flops between a and c production, and IV, black – asynchronous low c and intermediate a proteins expression. In addition, I introduced the next states: Ia, orange – for low b with intermediate and high c levels, as well as state IIa, yellow – for intermediate b and high c proteins expression. It should be remembered that synchronous and asynchronous FFBL are the common features of real genetic networks (Figure S3).

VI. Scalable design

5.1. Compositional mechanisms of modularity; interaction, communication. As I mentioned, instead of small gradual changes like point mutations, the mechanisms of compositional evolution combine interdependent genetic modules that have evolved previously in parallel. Examples of compositional mechanisms in Nature include recombination, hybridization, symbiotic encapsulation and horizontal gene transfer (HGT), as exhibited in the history of major evolutionary transitions. Modules must persist as identifiable units to be assembled into a Goldschmidt's 'hopeful monster' [37]. Different species are constantly exchanging genes, often with viruses as the messengers. Microbes can pass on fragments of DNA to each other during horizontal gene transfer. Even an entire pathway can be transferred if the respective genes are placed close to each other on the DNA sequence [38]. Furthermore, 'travelling' pathways exert their functions in cells with different genetic background and in changing environments.

Interspecies gene transfer also occurs (at an unknown rate) among more complex species, including humans. We demonstrated HGT in some eukaryotic species, such as mussels [39], fish [40], and rabbit [41] in lab conditions. Figure 9 shows the result of successful pcDNA3-*lacZ* sperm-mediated gene transfer into fish *M. fossilis*. In that case, a combination of electrical impulses with dimethyl sulfoxide (DMSO) treatment was used to improve the efficacy.

5.2. Design of complex systems: make parts, repeat them, and change them. It is well known that recursive functions generate fractals (Figure 2), less known are recursive functions in agents. Unfortunately, collective behavior and interaction between agents have been mostly ignored by biochemists and molecular biologists. Here, I give an example of a simple system that can produce a complex behavior.

An artificial world of Sperm Cells and Ova was investigated in the agent-based simulation [42]. If the meeting of a Spermatozoon and Ovum leads to a new Spermatozoon and new Ovum with a new genome, then 'genome mutations' will have occurred. This system demonstrates different kinds of behavior depending on the 'mutation' parameter R .

In detail, each creature has a circular genome consisting of 1024 'genes', only one of them is active and color coded with mod1024. The state of each creature is described by following recursive function:

$$\begin{aligned}T(i+1) &\leftarrow ([T(i) + P(i)] / 2 * R) \bmod 1024 \\P(i+1) &\leftarrow T(i+1),\end{aligned}$$

where $T(i)$ is the color code of the individual Spermatozoon and $P(i)$ is the color code of the individual Ovum at the time i of breeding. R is the mutation parameter on the interval $]0, 4]$.

5.3. Emergent behavior depending on mutation parameter. The system demonstrated ordered ($R \leq 1$) and complex ($R > 1$) regimes, such as (1) stable focus, $R=1$, (2) periodic, $R=1.01$, and (3) chaotic, $R=3$ regimes, as well as (4) strange attractor, $R=4$ (Figure 10). This complex and unexpected behavior of the artificial world of two agents – Sperm Cells and Ova – appeared from the collective dynamics of the distributed creatures and parallel execution of the recursion.

VII. Conclusion

Search for modularity in Nature is similar to pattern recognition, a native ability of the human mind. However, the discovery of hidden rules and algorithms leading to modularity is not an easy mission. The first attempt to describe evolutionary processes in terms of modules was carried out by John Holland in his building block hypothesis [43]. Richard Watson introduced three different algorithmic paradigms of evolution [8]. Watson classified systems on the basis of interdependency of variables. He considered weak, modular and arbitrary interdependencies of variables; smooth, spike and ruffle fitness landscapes; different optimization methods such as hill-climbing, divide-and-conquer decomposition, exhaustive and random search; different kinds of complexity on the basis of the number of variables and the number of values for each variable. Finally, Watson provided an evolutionary analogy for each class that is to say gradual evolution, compositional evolution, and impossible analogy or 'intelligent design'.

The author of this manuscript has seen the origin of modularity in a specific interaction between components of complex systems. *Exchange of modules has appeared as a formal origin of living entities.* It is my firm belief that interaction between a given system or multiple systems (agents) and an environment leads to the change of this environment along with a new adaptive behavior of its inhabitants. These kinds of interactions lead to a progressive increasing complexity of the system and environment as a result of more sophisticated interactions between components, see [44]. In this full of life situation resulting from the 'Red Queen Effect', a natural *de novo* design is preferable in the sense of economy than a possible reconstruction of 'old' creatures. I expect this imaginable picture is a possible response to the great question of why complex systems finally fail or die.

An astonishing observation on the global bacterial sulfur metabolism that the lateral gene transfer affected 16S rRNA phylogeny leads me to the following important conclusion: "the lateral gene transfer supports the modularization on the global scale" with the corollary: "the recombination provides modularization in protein structures". I have explained in this paper the model of the abstract *Argo-machine*, which is driven by 'oracle' words that are generated in turn by the environment. The model was inspired by data on Argonaute proteins and siRNA/RISC complex [45], also by ping-pong

amplification loop mechanism and the existence of transposon-rich piRNA genome clusters [46], by rampant horizontal gene transfer in prokaryotes [47, 48], as well as by my own experiments on sperm-mediated gene transfer [39-41], and the latest achievements in the field of SMGT [49, 50]. The view of molecular genetics and epigenetics mechanisms like the 'molecular computation' is in my opinion a very fruitful concept. A good example is research on developmental genome rearrangements in ciliates provided by Landweber and Kari [51]. A few simple rules and algorithms embedded in autonomous agents can lead to a broad variety of the system behaviour that is also demonstrated within this paper in terms of an agent-based approach. The problem of interaction between modules is very important in the practice of genetic engineering because of the desirable compatibility between modules in synthetic design. That was the reason to discuss xenologs versus orthologs in my research on sulfur metabolism in environmental bacterial populations [7].

Before I finish, I would like to summarize the key points:

- A module is the part which operates independently of other components in the system
- Functional modularity is independence in space and time
- Modularity is driven by the interaction and communication of components
- A set of modules can be joined in different ways when the environment changes (e.g. HGT)
- Origin of modularity is in the compositional evolution
- Modularity expands parallel development and enhances evolvability
- Specific interaction between modules is a subject of compositional design of complex systems
- Modularity is the relationship between the whole and the parts.

I believe that the understanding of evolution as a computational process of an ever changing environment can help us find design principles of biological systems. Modular systems consist of subsystems that work autonomously and exert specific functions. Biology can be described in terms of modules; furthermore modularity could be considered as a scientific issue, because modular and hierarchical structures demonstrate evolutionary benefits [52].

Acknowledgments

The author is solely responsible for any unconventional conclusions presented in this manuscript. In addition, I would like to express my gratitude to Bert Schnell, Heinz Eikmeyer, Mikhail Kats, Genaro Juarez Martinez, Vladik Avetisov, Irina Shchit, Irena Kuznetsova, Sergey Golutvin, Polina Tereshchuk, Svetlana Santer, Konrad Diwold, Martin Schneider, Andrew Phillips, Steven Benner, and many others who are not on the list.

References

- [1] Needham J. On the dissociability of the fundamental processes in ontogenesis // *Biological Reviews*. 1933; 8: 180-233.
- [2] Andrews J. Bacteria as modular organisms // *Annual Review of Microbiology*. 1998; 52: 105-26.
- [3] Ancel LW, Fontana W. Plasticity, evolvability, and modularity in RNA // *J Exp Zool*. 2000; 288(3): 242-83.
- [4] Bolker JA. Modularity in development and why it matters to Evo-Devo // *American Zoologist*. 2000; 40: 770-6.
- [5] Kashtan N, Alon U. Spontaneous evolution of modularity and network motifs // *Proc Natl Acad Sci U S A*. 2005; 102(39): 13773-8.
- [6] Wagner GP, Pavlicev M, Cheverud JM. The road to modularity // *Nat Rev Genet*. 2007; 8(12): 921-31.
- [7] Kuznetsov A. Modularity and distribution of sulfur metabolism genes in bacterial populations: search and design // *Journal of Computer Science & Systems Biology*. 2010; 3(5): 91-106.
- [8] Watson RA. *Compositional Evolution: The Impact of Sex, Symbiosis, and Modularity on the Gradualist Framework of Evolution* // Vienna Ser Theor Biol. A Bradford Book. 2006.
- [9] Guimerà R, Nunes Amaral LA. Functional cartography of complex metabolic networks // *Nature*. 2005; 433(7028): 895-900.
- [10] Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology // *Nature*. 1999; 402(6761 Suppl): C47-52.
- [11] Ihmels J, Friedlander G, Bergmann S, Sarig O, Ziv Y, Barkai N. Revealing modular organization in the yeast transcriptional network // *Nat Genet*. 2002; 31(4): 370-7.
- [12] Jayaswal V, Lutherborrow M, Ma DD, Yang YH. Identification of microRNA-mRNA modules using microarray data // *BMC Genomics*. 2011; 12: 138.
- [13] Liu B, Liu L, Tsykin A, Goodall GJ, Green JE, Zhu M, Kim CH, Li J. Identifying functional miRNA-mRNA regulatory modules with correspondence latent dirichlet allocation // *Bioinformatics*. 2011; 26(24): 3105-11.
- [14] Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles // *Proc Natl Acad Sci U S A*. 1999; 96(8): 4285-8.
- [15] Schuster S, Pfeiffer T, Moldenhauer F, Koch I, Dandekar T. Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae* // *Bioinformatics*. 2002; 18(2): 351-61.
- [16] Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data // *Nat Genet*. 2003; 34(2): 166-76.
- [17] Kuznetsov A. Evolution by communication: a revision of sperm-mediated gene transfer // *Frontiers in the Convergence of Bioscience and Information Technologies: FBIT 2007*, Ed. Daniel Howard et al.; Proceedings of the FBIT 2007 International Conference, Jeju Island, Korea, 11-13 October 2007, P. 322-4.

- [18] Callebaut W, Rasskin-Gutman D. (Eds.) Modularity: Understanding the development and evolution of natural complex systems // Vienna Ser Theor Biol. MIT Press. 2005.
- [19] Kuznetsov AV. Study of physical, chemical, and immunological properties of the glial fibrillary acidic protein // Graduate Thesis. Dnipropetrovs'k University. 1983.
- [20] Kuznetsova IV. Glial fibrillary acidic protein of the human brain // Graduate Thesis. Dnipropetrovs'k University. 1984.
- [21] Kuznetsov A. Assembling by adhesion rules at the nanoscale // DECOI 2007: Design of Collective Intelligence, International Summer School on Collective Intelligence and Evolution, Amsterdam, Holland, 20-24 August 2007.
- [22] Simon HA, Ando A. Aggregation of Variables in Dynamic Systems // *Econometrica*. 1961; 29(2): 111-38.
- [23] Kuznetsov A. Barbie nanoatelier // *IET Synthetic Biology*. 2007; 1(1-2): 7-12.
- [24] Kuznetsov A. Magnetic properties of endohedral complexes $\text{Co}_5@C_n$ depending upon the size and symmetry of fullerenes as well as orientation of cobalt cluster // *Computational Materials Science*. 2012; 54: 204-7.
- [25] Kuznetsov A. From carbides to Co_5 and Co_{13} metallofullerenes: first-principles study and design // *Am J Biomed Eng*. 2012; 2(1): 32-8.
- [26] Strähle U, Blader P. The basic helix-loop-helix proteins in vertebrate and invertebrate neurogenesis // in *Modularity and Evolution*. Eds. Gerhard Schlosser and Günter P. Wagner. University of Chicago Press. 2004.
- [27] Alon U. An Introduction to Systems Biology: Design Principles of Biological Circuits // *Chapman & Hall/CRC Mathematical & Computational Biology*. 2006.
- [28] Koonin EV, Martin W. On the origin of genomes and cells within inorganic compartments // *Trends Genet*. 2005; 21: 647-54.
- [29] Wächtershäuser G. Before enzymes and templates: theory of surface metabolism // *Microbiol Rev*. 1988; 52: 452-84.
- [30] Kuznetsov A, Schmitz M, Mueller K. On bio-design of Argo-machine // *Explorations in the complexity of possible life: abstracting and synthesizing the principles of living systems*, Ed. by Peter Dittrich, Stefan Artmann; *Proceeding of the 7th German Workshop on Artificial Life (GWAL-7)*, Jena, Germany, 26-28 July 2006. P. 125-33.
- [31] Wolfram S. *A New Kind of Science* // Wolfram Media, Inc. 2002.
- [32] Milner R. Functions as Processes // *Mathematical Structures in Computer Science*. 1992; 2(2): 119-41.
- [33] Priami C. Stochastic pi-Calculus // *Comput. J*. 1995; 38(7): 578-89.
- [34] Phillips A. The Stochastic Pi-Machine // Available from <http://research.microsoft.com/~aphillip/spim/> (2007).
- [35] Gillespie D. Exact stochastic simulation of coupled chemical reactions // *J Phys. Chem*. 1977; 81: 2340-61.
- [36] Kuznetsov A. Genetic networks described in stochastic Pi Machine (SPiM) programming language: compositional design // *Journal of Computer Science & Systems Biology*. 2009; 2(5): 272-82.
- [37] Goldschmidt RB. Evolution as viewed by one geneticist // *American Scientist*. 1952; 40: 84-98.
- [38] Wren B. Microbial genome analysis: insights into virulence, host adaptation and evolution // *Nat Rev Genet*. 2000; 1: 30-9.
- [39] Kuznetsov AV, Pirkova AV, Dvorianchikov GA, Panfertsev EA, Gavriushkin AV, Kuznetsova IV, Erokhin BE. Study of the transfer of foreign genes into mussel *Mytilus Galloprovincialis* Lam. eggs by spermatozoa // *Ontogenez*. 2001; 32(4): 309-18.
- [40] Andreeva LE, Sleptsova LA, Grigorenko AP, Gavriushkin AV, Kuznetsov AV. Loach spermatozoa transfer foreign DNA, which expression is discovered in the early development stages // *Genetika*. 2003; 39(6): 758-61.
- [41] Kuznetsov AV, Kuznetsova IV, Schit IYu. DNA interaction with rabbit sperm cells and its transfer into ova in vitro and in vivo // *Molecular Reproduction & Development*. 2000; 56(2): 292-7.
- [42] Kouznetsov AV. Toy SMGT // *Alife mutants hacking session on systems and organisms (AMHSO)*, Rule 110 Winter Workshop, Bielefeld, Germany, 6-13 March 2004.
- [43] Holland JH. *Adaptation in Natural and Artificial Systems* // Ann Arbor, MI: The University of Michigan Press. 1975.
- [44] Szöllosi GJ, Derényi I. Congruent evolution of genetic and environmental robustness in micro-RNA // *Mol Biol Evol*. 2009; 26(4): 867-74.
- [45] Parker JS, Roe SM, Barford D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex // *Nature*. 2005; 434(7033): 663-6.
- [46] Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race // *Science*. 2007; 318(5851): 761-4.
- [47] Boto L. Horizontal gene transfer in evolution: facts and challenges // *Proc Biol Sci*. 2010; 277(1683): 819-27.
- [48] Omelchenko MV, Makarova KS, Wolf YI, Rogozin IB, Koonin EV. Evolution of mosaic operons by horizontal gene transfer and gene displacement in situ // *Genome Biol*. 2003; 4(9): R55.
- [49] Sciamanna I, Vitullo P, Curatolo A, Spadafora C. Retrotransposons, reverse transcriptase and the genesis of new genetic information // *Gene*. 2009; 448(2): 180-6.
- [50] Spadafora C. A reverse transcriptase-dependent mechanism plays central roles in fundamental biological processes // *Syst Biol Reprod Med*. 2008; 54(1): 11-21.
- [51] Landweber LF, Kari L. The evolution of cellular computing: nature's solution to a computational problem // *Biosystems*. 1999; 52(1-3): 3-13.
- [52] Simon HA. The architecture of complexity // *Proceedings of the American Philosophical Society*. 1962; 106(6): 467-82.

Appendix

```
(* Repressilator *)
directive sample 50000.0
directive plot !a as "a"; !b as "b"; !c as "c"
directive graph

val bind = 10.0      (* protein binding - r *)
val transcribe = 0.1 (* constitutive expression - epsilon *)
val unblock = 0.001 (* repression delay - eta *)
val degrade = 0.001 (* protein decay - delta *)

(* transcription factor *)
let tr(p:chan()) =
  do !p; tr(p)
  or delay@degrade

(* neg gate *)
let neg(a:chan(), b:chan()) =
  do ?a; delay@unblock; neg(a,b)
  or delay@transcribe; (tr(b) | neg(a,b))

(* circuit *)
new a@bind:chan()
new b@bind:chan()
new c@bind:chan()

run (neg(a,b) | neg(b,c) | neg(c,a))
```

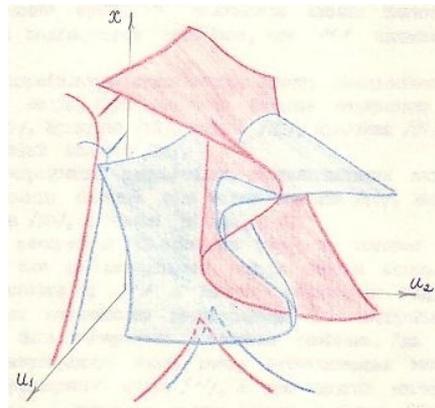


Figure 1. Soluble (red) and insoluble (blue) forms of GFA protein in nervous cells, where are x – GFA protein content, u_1 – malignancy, u_2 – putative immune response [20].

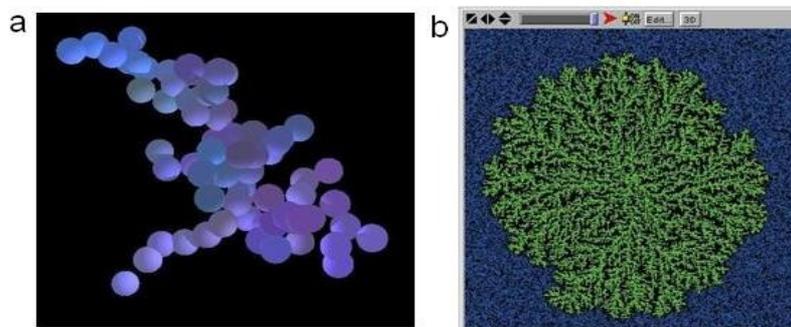


Figure 2. Diffusion limited aggregation (DLA), with a – 3D simulation, b – 2D simulation.

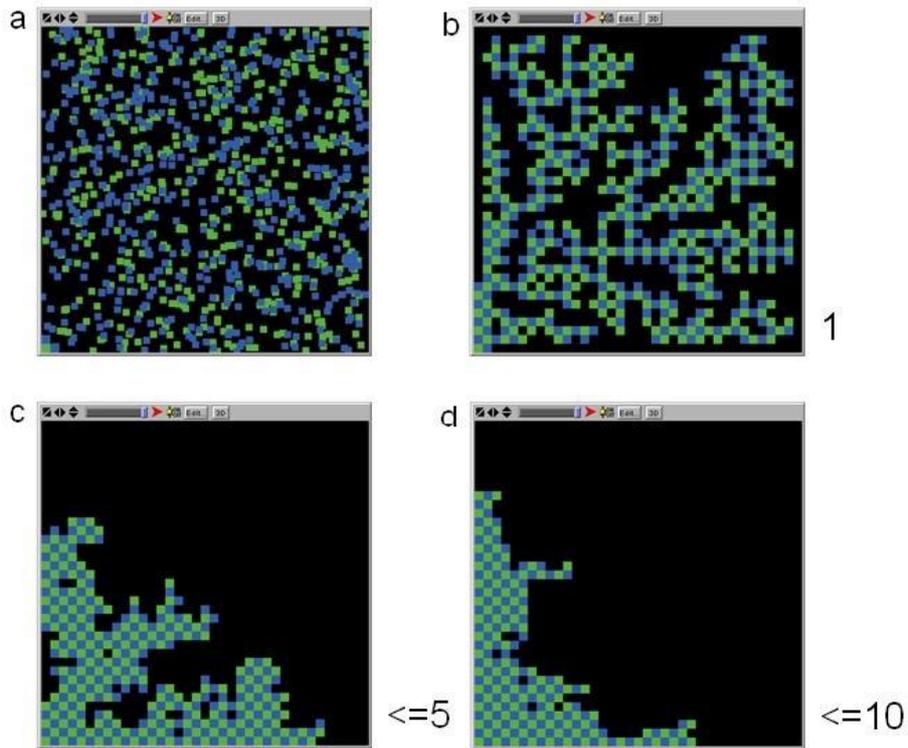


Figure 3. The chessboard pattern formation by adhesion rules in presence of DLA, where are a – initial state, b-d – final states with different values of diffusion.

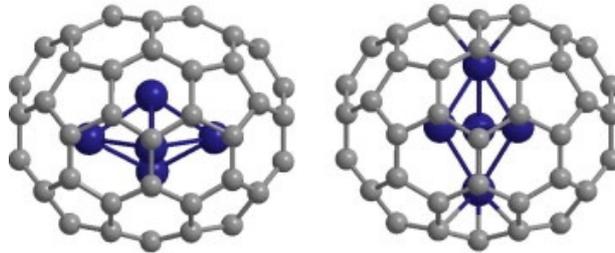


Figure 4. Cobalt-5 cluster within fullerene C70 in along and across orientations.

AprA	199	A	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1432856135	7	G	RRSIPPTSSVPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1433546210	118	A	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1433688352	62	C	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1433708223	170	A	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1433766319	1	--	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1433547459	1	--	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE
1434025767	1	--	GGASNIIFPRSVGEGAGRWVYAPWSSGSAYGLIMIGAGAKMTQENRIVLARPKDGYGPGVAYFLHLKTYTQNCNGEBEYBSKWFPGLOKRVGKEYLDPE

AprA	298	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1432856135	107	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1433546210	217	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1433688352	154	A	APYCHAAITPTCLR-NHALINEINAGRCPITHMVTMEALAALGETMDKKELEHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1433708223	269	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1433766319	99	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1433547459	1	--	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG
1434025767	17	A	ASHLTHRFIPTCLR-NHALINEINAGRCPITHMVTMEA-----EQDPHLEBEIGWVNFGLMTVGOAVLWAATDVNPKYENPELTTSEPYWVGSHATG

Figure 5. Protein alignment of the DNA reads from microorganisms living in the gutless worm *O.algarvensis* to the AprA polypeptide from bacterium *V.okutanii*, where are AprA – protein from *V.okutanii*, numbers – reads from the Gutless Worm database, yellow – sufficient insertions, deletions and exchanges.

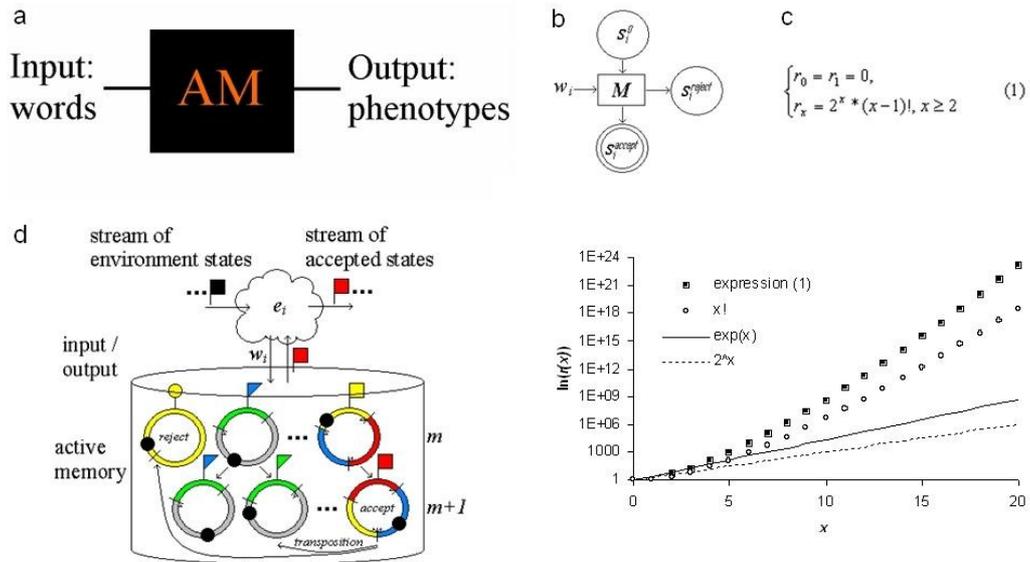


Figure 6. Schematics of Argo-machine, with a – block diagram, b – state diagram, c – combinatorial analysis, d – functional diagram.

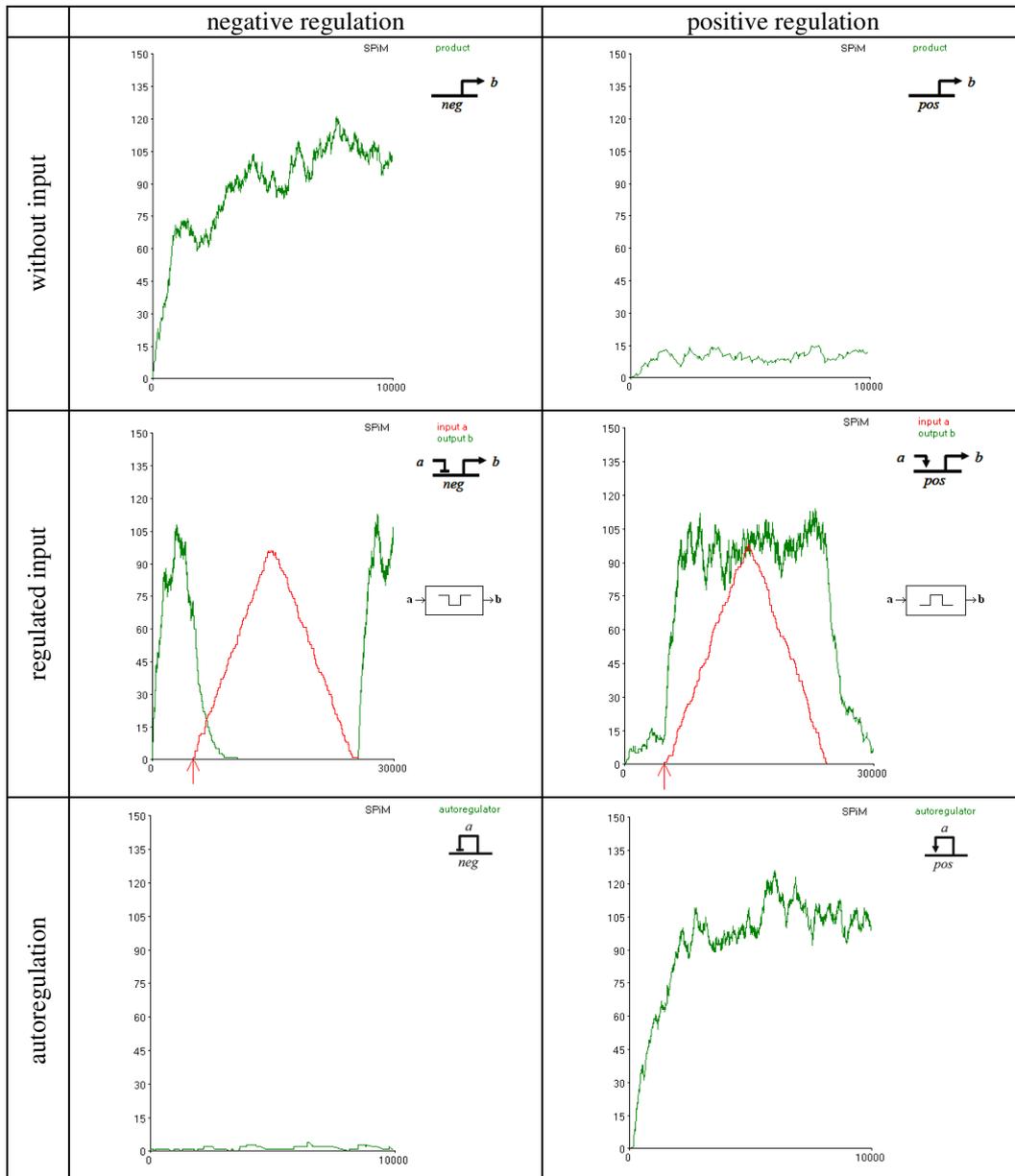


Figure 7. Basic genetic gates.

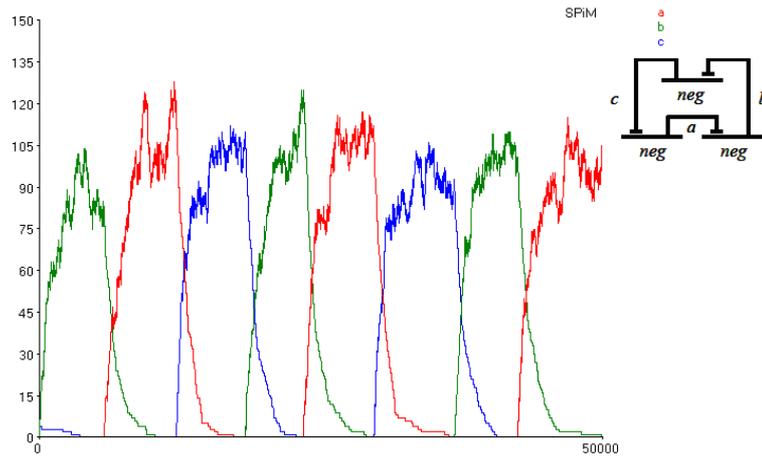


Figure 8. Repressilator, where $r = 10.0$, $\delta = 0.001$, $\varepsilon_n = 0.1$, $\eta_n = 0.001$ [36].



Figure 9. Sperm-mediated gene transfer of the plasmid pcDNA3-lacZ into *M.fossilis*.

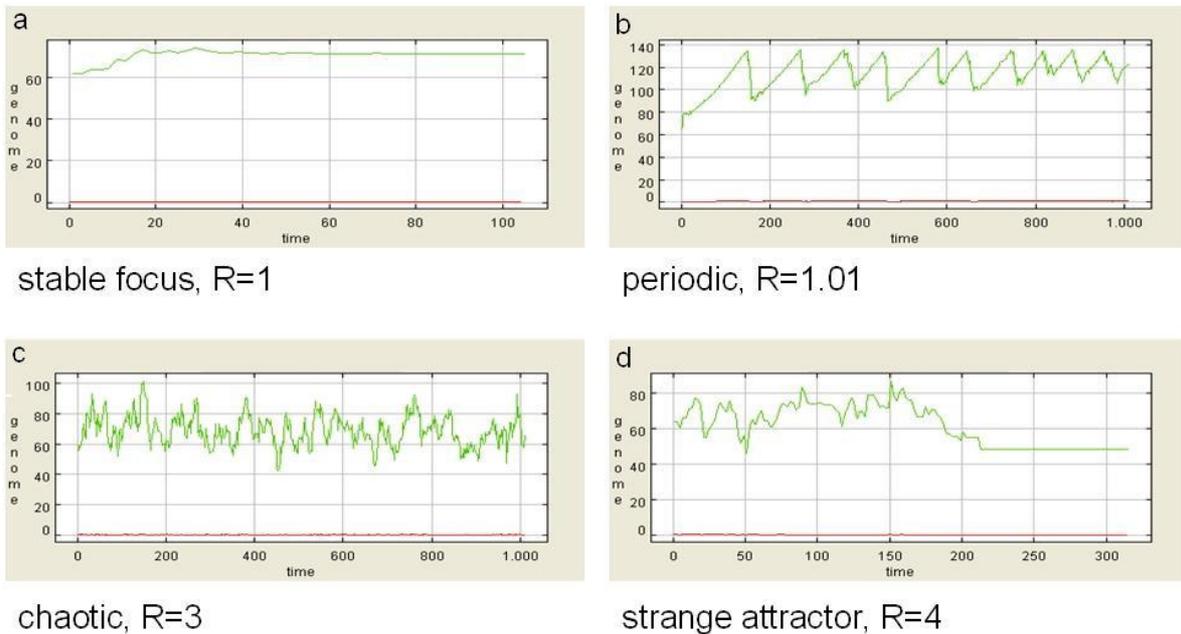


Figure 10. Behavior depending on mutation parameter R , where the abscissa is a time of simulation, the ordinate is an average genome.

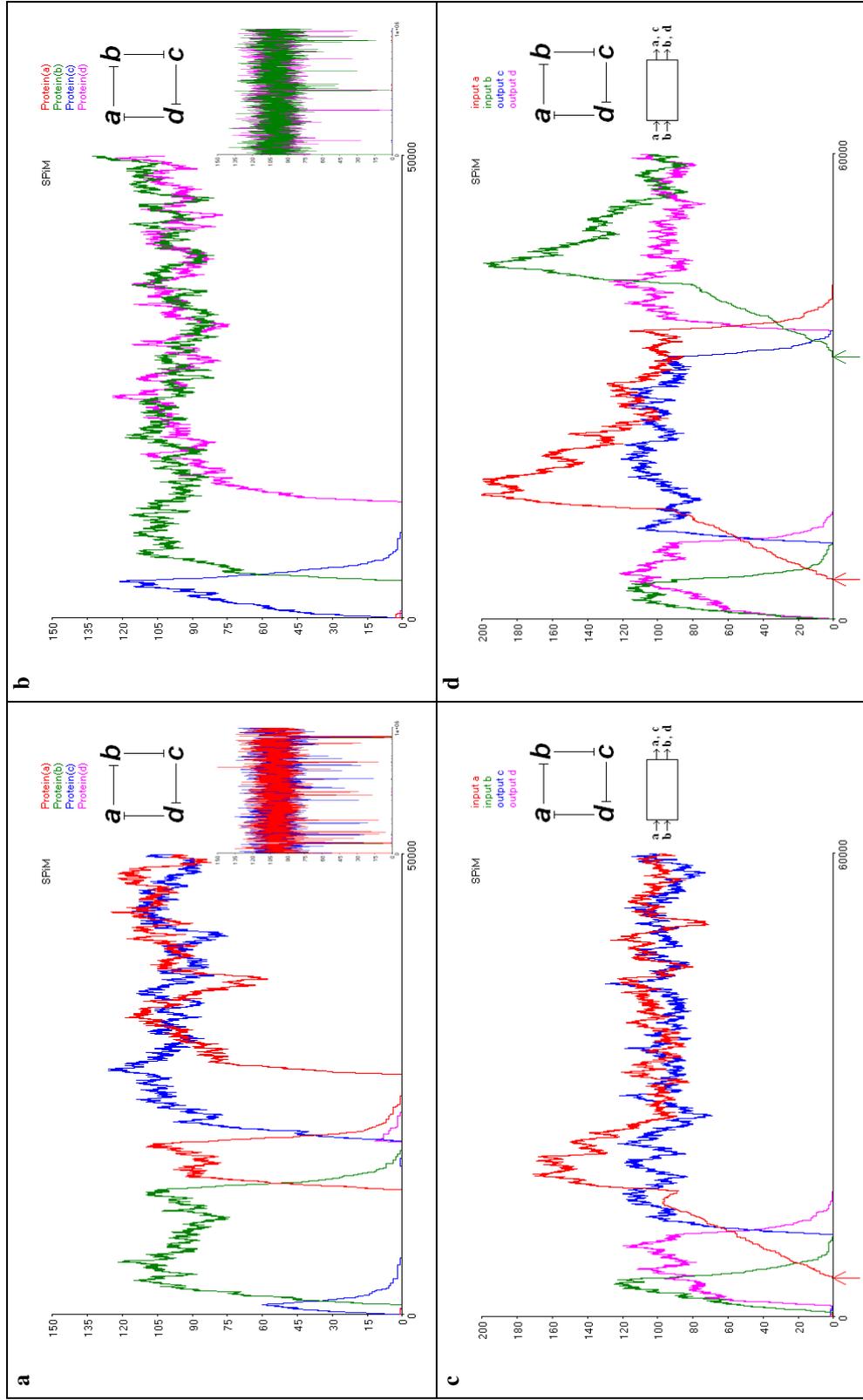


Figure S1. Memory, where a – initial *ac* state, b – initial *bd* state, c – input *a* switches *bd* to *ac* state, d – input *b* switches *ac* back to *bd* state.

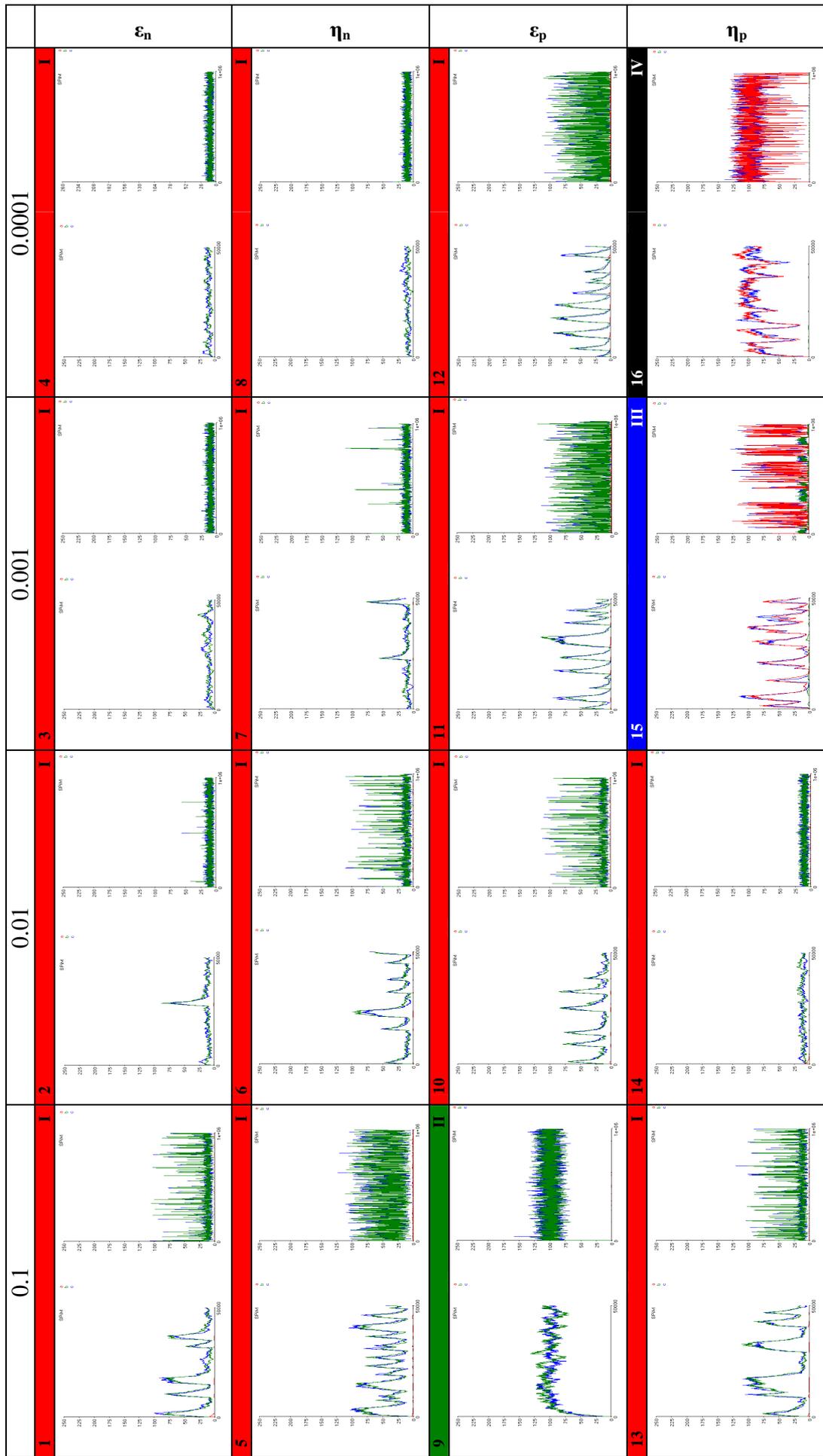


Figure S2. Coherent FFBL (short- and long-time simulations), see [36] for details.

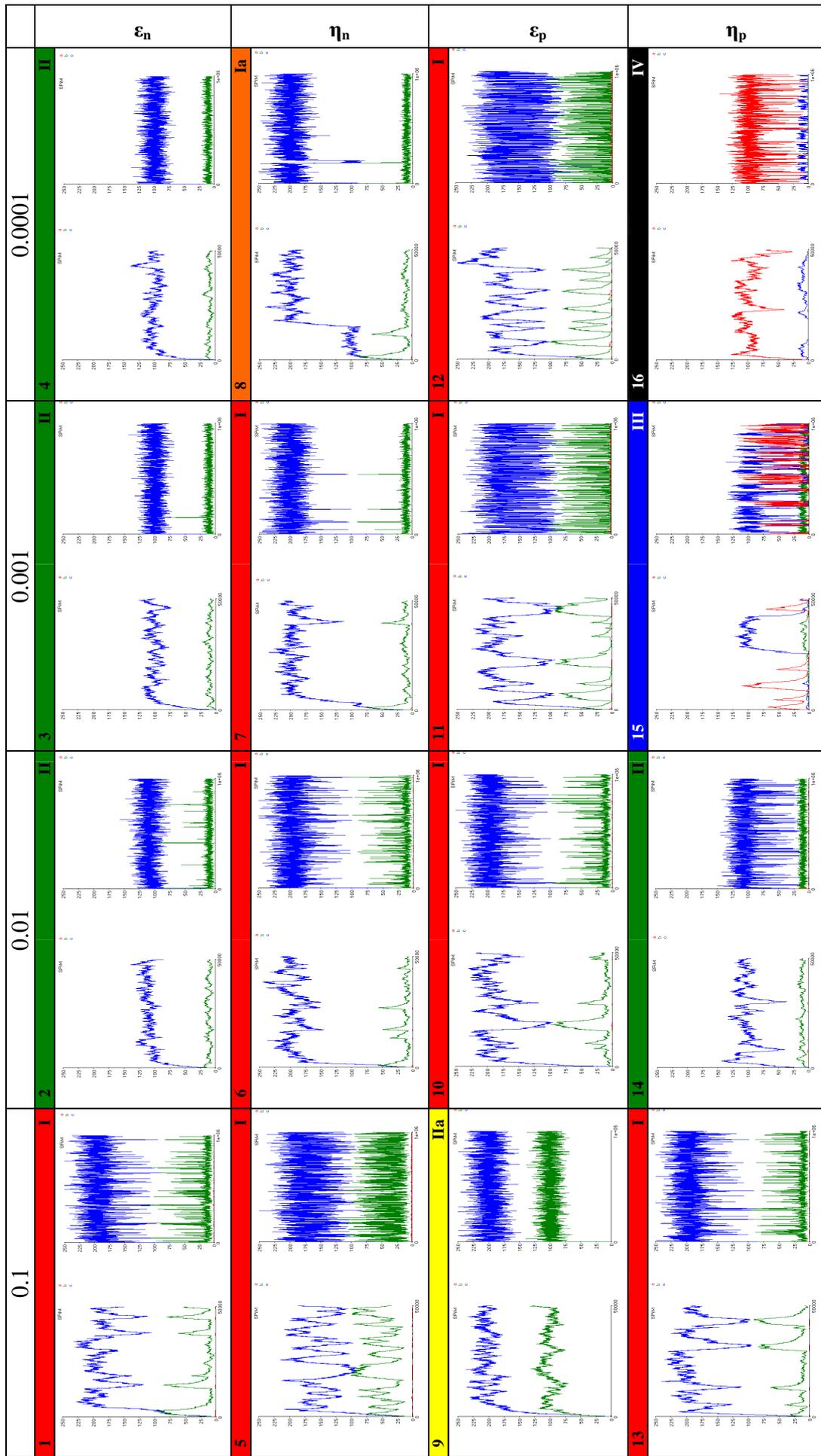


Figure S3. Incoherent FFBL (short- and long-time simulations), see [36] for details.

A New Approach for Disparity Map Estimation from Stereo Image Sequences using Hybrid Segmentation Algorithm

Patrik Kamencay¹, Martina Zachariasova², Martin Breznan³, Roman Jarina⁴,
Robert Hudec⁵, Miroslav Benco⁶, Slavomir Matuska⁷

^{1, 2, 3, 4, 5, 6, 7}Department of Telecommunications and Multimedia, University of Zilina, Slovakia

ABSTRACT: In this paper, a stereo matching algorithm based on image segments is presented. We propose the hybrid segmentation algorithm that is based on a combination of the Belief Propagation and K-Means algorithms with aim to refine the final sparse disparity map by using a stereo pair of images. Firstly, a color based segmentation method is applied for segmenting the left image of the input stereo pair (reference image) into regions. The main aim of the segmentation is to simplify representation of the image into the form that is easier to analyze and is able to locate objects in images. Secondly, results of the segmentation are used as an input of the SIFT-SAD matching method to determine the disparity estimate of each image pixel. This matching algorithm is proposed by combining Scale Invariant Feature Transform (SIFT) with the Sum of Absolute Difference (SAD). Finally, the comparisons between the three robust feature detection methods: Scale Invariant Feature Transform (SIFT), Affine SIFT (ASIFT) and Speeded Up Robust Features (SURF) are presented. The obtained experimental results demonstrate that the performance of our method is competitive and the final disparity maps are close to the ground truth data.

Keywords: Belief Propagation, K-Means, SIFT-SAD, hybrid segmentation, depth map

I. INTRODUCTION

This paper describes a set of algorithms for structure, motion automatic recovery and visualization of a 3D image from a sequence of 2D images. The important step to perform this goal is matching of corresponding pixels in the different views to estimate the disparity map and final depth map. The disparity of an image pixel is the distance of the corresponding world point from the camera center. Detecting objects, estimating their pose, geometric properties and recovering 3D shape information are a critical problem in many vision and stereo computer vision application domains such as robotics applications, high level visual scene understanding, activity recognition, and object modeling [1].

The structure and motion recovery system follows a natural progression, comprising the following phases:

- Image acquisition,
- Feature matching using SIFT descriptor,
- Image segmentation,
- feature detection using SIFT-SAD algorithm,
- Stereo geometry and image rectification,
- Estimation of disparity and final depth map.

A classical problem of stereo computer vision is the extraction of 3D information from stereo views of a scene. To solve this problem, knowledge of view properties and

feature point between views is needed. However, finding these points is notoriously hard to do for natural scenes. The fundamental idea behind stereo computer vision is the difference in position of a unique 3D point in two different images. As the object moves closer to the cameras, the relative position of object will change, and the positions in each image will move away from each other. In this way, is possible to calculate the distance of an object, by calculating its relative positioning in the two images. This distance between the same objects in two images is known as disparity [1]. Disparity map computation is one of the key problems in 3D computer vision.

This paper employed a new feature projection approach based on SIFT-SAD method using hybrid segmentation algorithm. A comparison between these two different approaches for the image segmentation (K-Means and Belief Propagation) is described in [2], [3].

The outline of the paper is as follows. In the next section, an overview of image segmentation methods (K-Means, Belief Propagation and hybrid segmentation) is introduced. In section III, the process of disparity map calculation from corresponding points using stereo matching algorithms is presented. The evaluation criterion used for comparing image segmentation algorithms is described in Section IV. Finally the experiment results and architecture of reconstruction algorithm are introduced in Section V, and brief summary is discussed in Section VI.

II. IMAGE SEGMENTATION

The main goal of the image segmentation is split the entire image into set of segments that cover image. In this chapter, K-Means, Belief Propagation and proposed hybrid segmentation algorithms will be presented.

A. K-Means segmentation

K-means algorithm is statistical clustering algorithm. Data clustering is method that creates groups of objects (clusters). K-means algorithm is based upon the index of similarity or dissimilarity between pairs of data components [3]. This type of algorithm is popular for simplicity, implementation and it is commonly used for grouping pixels in images. But this algorithm has three basic disadvantages [3], [5]:

- K , the number of clusters must be determined.
- Different initial conditions produce different results.
- The data far away from center pull the centers away from optimum location.

Let $X = \{x_1, x_2, \dots, x_n\} \subset R^p$ be a finite set of data where N is the number of data items and R^p is p -dimensional Euclidean space [5]. Let V_{KN} be the set of matrices ($K \times N$,

$2 \leq K < N$), where K is number of clusters. A K partition of X is defined:

$$M_K = \left\{ U \in V_{K,N} \mid \begin{array}{l} u_{ik} \in \{0,1\}, \forall i,k; \\ \sum_{i=1}^K u_{ik} = 1, \forall k \end{array} \right\}, \quad (1)$$

where $u_{ik} = 1$ denotes that component x_k belongs to cluster i , $u_{jk} = 0$ denotes that component x_k is out of the cluster j . The objective function J_K is:

$$J_K(U, K) = \sum_{i=1}^K \sum_{k=1}^N u_{ik} d_{ik}^2, \quad (2)$$

where d_{ik} is Euclidean distance between component x_k and cluster v_i :

$$d_{ik} = \|x_k - v_i\| \quad (3)$$

Optimal K clusters of X is produced by minimization of objective function $J_K(U, V)$ [3], [5].

B. Belief Propagation

Belief Propagation is an iterative inference algorithm for graphical models such as MRF which is based on a message passing principle that propagates messages in the network [2], [4]. The above model contains only pairwise cliques, and the joint probability over the 3D volume is

$$P = \prod_{i \neq j} \psi_{i,j}(s_i, s_j) \prod_k \phi_k(s_k, d_k), \quad (4)$$

where s_j and d_j represent state node and data node separately [2], [4]. ψ is the state transition function between a pair of different hidden state nodes and ϕ is the measurement function between the hidden state node and observed data node. N represents the total number of state or data nodes in the 3D volume. Under the squared loss function, the best estimate for node s_j is the mean of the posterior marginal probability (minimum mean squared error estimate, MSE estimate):

$$s_{jMSE} = \sum_{s_j} s_j \sum_{s_i, i \neq j} P(s_1, \dots, s_N, d_1, \dots, d_N), \quad (5)$$

where the inner sum gives the marginal distribution of s_j [1], [2], [4].

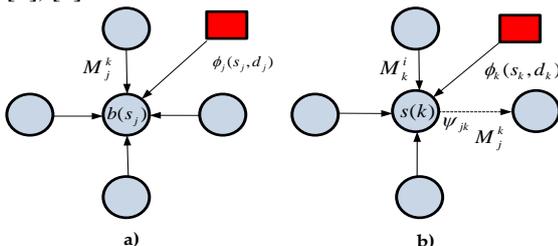


Fig. 1. (a) Computing belief, (b) Computing message [4]

Since the joint probability involves all the hidden state nodes and data nodes in the 3D volume, it is hard to compute the MSE estimate based on the implicit

multivariable probability distribution. However belief propagation messages are effective to compute the MSE estimate recursively. Each hidden state node has a belief, which is a probability distribution defining the node's motion likelihood. Thus the MSE estimate of one node is computed as:

$$s_{jMSE} = \sum_{s_j} s_j b(s_j), \quad (6)$$

where

$$b(s_j) = \phi_j(s_j, d_j) \prod_{k \in Neighbor(j)} M_j^k, \quad (7)$$

is the belief at node s_j and k runs over all neighboring hidden state nodes of node s_j . The belief at node s_j is the product of all the incoming messages M and the local observed data message ($\phi_j(s_j, d_j)$). The computation is shown in Fig. 1 (a). The passed messages specify what distribution each node thinks its neighbors should have. Fig. 1 (b) shows how to compute the message from node s_k to s_j :

$$M_j^k = \sum_{s_k} \psi_{jk}(s_j, s_k) b(s_k). \quad (8)$$

After multiplying all the incoming messages M from neighboring nodes (except from the node s_j) and the observed data message ($\phi_k(s_k, d_k)$), the product is evolved from the message-sender to the message-receiver by transition function $\psi_{jk}(s_j, s_k)$ [2], [4].

C. Hybrid segmentation algorithm

Hybrid methods are created by combining two or more image segmentation algorithms. In our image analysis, a hybrid algorithm, which is produced by the combination Belief Propagation [2], [4] and K-Means [3], [5] is used. This approach brings together the advantages of both segmentation algorithms. K-Means is quick and Belief Propagation is very accurate segmentation. Diagram of hybrid segmentation algorithm is shown in Fig. 2.

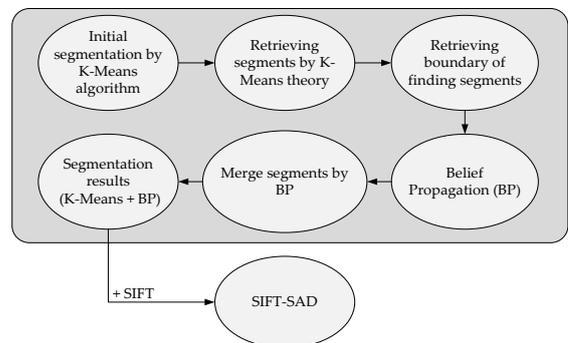


Fig. 2. Hybrid segmentation algorithm

First, we apply image filtering by Mean Shift algorithm. This step is very useful for noise removing, smoothing and image segmentation. For each pixel of an image, the set of neighboring pixels is determined. Let X_i be the input and Y_i filtered image, where $i = 1, 2, \dots, n$. The filtering algorithm comprises of the following steps [16]:

- Initialize $j = 1$ and $y_{i,1} = p_i$.
- Compute through the Mean Shift the mode where the pixel converges.
- Store the component of the gray level of the calculated value $Z_i = (x_i, y_{i,c})$ at Z_i , where x_i is the spatial component and $y_{i,c}$ is the range component.

Secondly, the image is split into segments using K-Means algorithm. In the third step, means of segments are retrieved by applying K-Means theory. Fourth, the small segments are merged together to the most similar adjacent segments by the Belief propagation method [10]. Finally, we have integrated our proposed hybrid segmentation algorithm with the Sum-of-Absolute-Differences (SAD) stereo matching algorithm and SIFT descriptor. This hybrid SIFT-SAD method (HSAD) is able to produce highly accurate depth map.

III. FEATURE CALCULATION

The Feature matching algorithm improves precision of disparity calculation. This kind of algorithms extracts object's suitable features in the 3D scene, for example, segments of edges or contours in the left and right stereo images [8]. In the following stage, disparity map is calculated from corresponding points of the features.

A reconstruction of the disparity map from the left and right stereo pair is known as the stereo matching algorithm. The detection feature points must be matched. There exist several matching techniques based on various algorithms, e.g. Correlation (C), Normalized Cross Correlation (NCC), Sum of Squared Differences (SSD) and Sum of Absolute Differences (SAD) algorithms. In our case we used SAD matching algorithm [6].

A. SIFT descriptor

Scale Invariant Feature Transform (SIFT) is a local descriptor of image features insensitive to illuminant and other variants that is usually used as sparse feature representation [7]. SIFT features are features extracted from images to help in reliable matching between different views of the same object [6]. The extracted features are invariant to scale and orientation, and are highly distinctive of the image. They are extracted in four steps. The first step computes the locations of potential interest points in the image by detecting the maxima and minima of a set of Difference of Gaussian (DoG) filters applied at different scales all over the image. Then, these locations are refined by discarding points of low contrast. An orientation is then assigned to each key point based on local image features. Finally, a local feature descriptor is computed at each key point. This descriptor is based on the local image gradient, transformed according to the orientation of the key point to provide orientation invariance. Every feature is a vector of dimension 128 distinctively identifying the neighbourhood around the key point [7].

B. Sum of Absolute Differences (SAD)

The SAD is widely used metric for block matching in stereo images. It works by taking the absolute value of the difference between each pixel in the original block and the corresponding pixel in the block being used for comparison. These differences are summed over the block to create a simple metric of block similarity, the $L1$ norm of the

difference image [8]. The cost function $C(p)$ on the basis of SAD is computed as follows:

$$C(p) = \sum_{(p') \in W(p)}^N |I_L(p') - I_R(p')|, \quad (9)$$

where p is the reference pixel at which the SAD is computed and p' represent the pixels belonging to the neighborhood of the pixel p . $W(p)$ is square window that surrounds the position (p_x, p_y) of the pixel. The minimum difference value over the frame indicates the best matching pixel, and position of the minimum defines the disparity of the actual pixel [6], [8].

Quality of 3D disparity map depends on square window size, because a bigger window size corresponds to a greater probability of correct pixel disparity calculated from matched points, although the calculation gets slower [19].

C. SIFT-SAD algorithm

The final performance of stereo matching algorithms depends on the choice of matching cost. In our experiment we proposed SIFT-SAD matching method as matching cost. SIFT descriptor delivers most of local gradient information and SAD provides local intensity information. SIFT-SAD consists of two parts. Firstly, we get the $L1$ distance of SIFT between pixel p in the left image and $p+d_p$ in the right image [6], [8].

$$D_{SIFT}(d_p) = \|x_L(p) - x_R(p+d_p)\|, \quad (10)$$

where d_p is the disparity of pixel, $\|x_L(p) - x_R(p+d_p)\|$ is the $L1$ distance. Next, we define SAD matching cost as:

$$D_{SAD}(d_p) = \exp(-SAD(p, p+d_p)), \quad (11)$$

where $SAD(p, p+d_p)$ is the SAD score in a square neighborhood searching window. Our algorithm computes the disparity for all pixels with window size dimension at square of 9×9 pixels. Finally, we use one dimensional Gaussian weight with a scale factor s to get the matching cost. The underlying assumption is that if a minimum corresponds to the true surface, the neighboring pixels should have near values at a similar depth [8].

IV. FIGURES EVALUATION CRITERION

With the increase in the number of developed algorithms for image segmentation, evaluation criterion for studying of segmentation is required.

The criterion used for comparing image segmentation algorithms presented in this article, is based on computing precision, recall and $F1$. These three parameters determine the algorithms efficiency by comparing boundaries their segments. Each of the algorithms is compared with segmentation by a human. Based on this comparison, precision, recall and $F1$ are computed. The definition of precision, recall and $F1$ is given by:

$$P = \frac{C}{C+F} \cdot 100\%, \quad (12)$$

$$R = \frac{C}{C + M} \cdot 100\%, \quad (13)$$

where C is the number of correct detected pixels that belongs to boundary, F is the number of false detected pixels and M is the number of not detected pixels.

$F1$ is combined measure from precision and recall. It is in high values if both precision and recall have high values and on the other hand, if one of them has low value, the value of the $F1$ is going down. The definition of $F1$ is given by:

$$F1 = \frac{2PR}{P + R} \cdot 100\%. \quad (14)$$

V. EXPERIMENTAL RESULTS

In this section, some of the obtained experimental results rectifying, matching points and generating sparse disparity map will be presented. The proposed architecture (see Fig. 3) has been tested on two input real images. This proposed algorithm based on the combination of SAD stereo matching algorithm with SIFT descriptor is faster, since a small portion of whole left and right images pixels are used for matching. In this experiment, we segment the reference image (in our case, the left image) using hybrid segmentation method. Then, for each segment we look at the associated pixel disparities.

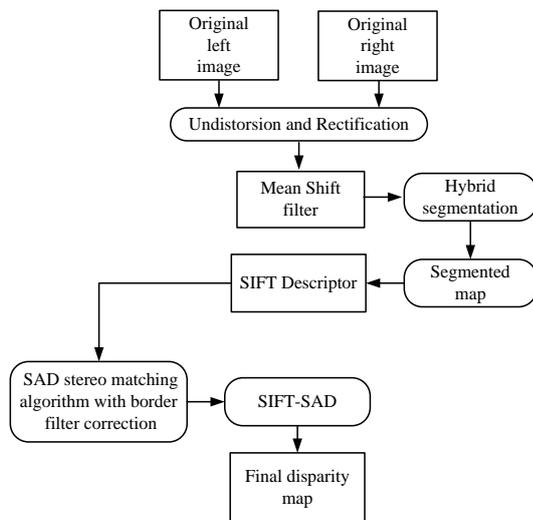


Fig. 3. Block diagram of proposed approach using SIFT-SAD algorithm

First, the edges are extracted using Harris method due to its good performance. Harris corner detector is a suitable starting point for the computation of positions of scale [12]. Next step is image rectification. It is transformation which makes pairs of conjugate epipolar lines become collinear and parallel to the horizontal axis (baseline). For the epipolar rectified images pair, each point in the left image lies on the same horizontal scan line as in the right image. This approach is used to reduce a search space for disparity map estimation algorithm. Then, we apply image filtering by Mean Shift filter. This step is very useful for noise removing, smoothing and image segmentation [13], [14].

After filtration, the filtered image is split into segments using K-Means segmentation algorithm. In the next step, the small segments are merged together to the most similar adjacent segments. Next, matching is performed, where a sparse disparity map is obtained. The match points can be obtained using SAD approach along the epipolar line [15]. The disparity map codifies the distance between the object and the camera - closer points will have maximal disparity and farther points will get zero disparity [16], [17]. Finally, we have integrated the segmentation algorithm K-Means with the SIFT-SAD method. This method consists of two parts: SIFT part and SAD part. As already stated, our experimental results proved that the SIFT descriptor is a very robust and reliable representation for the local neighborhood of an image point. This proposed approach is able to produce highly accurate disparity map.

The rectified left and right images are shown in Fig. 4. Both images were segmented by the proposed hybrid segmentation algorithm.



Fig. 4. Correspondence using window – based matching (rectified images)

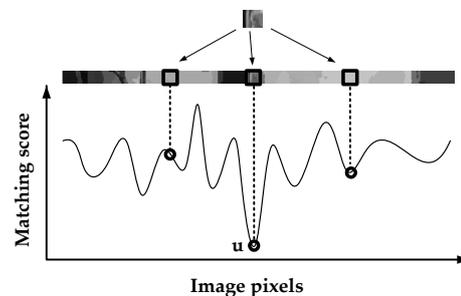


Fig. 5. Best matching pixel (minimum value over the row)

All disparity map pixels of all the segments were obtained using SIFT-SAD method along the same epipolar lines of the stereo images as it is shown in Fig. 5. The minimum value over the row in the right image is chosen to be the best matching pixel (see Fig. 5). The disparity is then calculated as the actual horizontal pixel difference [18].

The result of stereo matching process is a grayscale disparity map that indicates the disparity for every pixel with corresponding intensity. Quality of disparity map is represented as percentage of pixels with disparity errors (bad matching pixels) [17]:

$$P = \frac{1}{X * Y} \sum_{i=1}^X \sum_{j=1}^Y (|d_C(i, j) - d_T(i, j)|), \quad (15)$$

where $X * Y$ represent the size of the image, d_C is the computed disparity map of the test image and d_T is the truth disparity map [10].

$$d_T = \frac{fBI_{RES}}{D_T h}, \quad (16)$$

where D_T is ground truth depth map, h is height from the ground plane, $D_T * h$ is ground truth distance, B is baseline between the cameras, I_{RES} is image resolution and f is focal length.

Table 1 shows summary of overall performance. We compared performances obtained by the proposed method SIFT-SAD with those obtained by three common algorithms (SIFT, ASIFT and SURF - Speeded Up Robust features) [2], [6], [9]. Approximately 92 percent of the disparity values were found correctly for our proposed algorithm. The final disparity map is labeled as correct if it is within one pixel of the correct disparity. The ground truth disparity map [17] is the inverse of the ground truth distance scale by the image resolution and the focal length [15], [18]. Equation (16) shows how to calculate the ground truth disparity map from the depth map. The depth map is a 16 bit map with values ranging from 0 to 1 where the ground plane was at $D = 1$ and the cameras were at $D = 0$. D is distance of object from the camera.

TABLE 1. THE PERCENTAGE OF DISPARITIES FOUND CORRECTLY, DISPARITY ERROR AND THE DETECTED OCCLUSION THAT ARE CORRECT

	SIFT	ASIFT	SURF	SIFT-SAD
Disparity Correct [%]	86.69	82.07	89.78	92.35
Disparity Error [%]	13.31	17.93	10.22	7.65
Occlusion Correct [%]	67.45	65.32	72.76	72.03

First, three segmentation algorithms were compared using automatic algorithm evaluating the precision of segmentation, as is shown in Table 2. This plays important role for two reasons: (1) it can be placed into a feedback loop to enforce another run of segmentation algorithm that may include more sophisticated steps for high precision segmentation and (2) the outcome of this evaluation can be treated as a quality factor and thus can be used to design a quality driven adaptive recognition system.

TABLE 2. BEST RESULTS OF IMAGE SEGMENTATION ALGORITHM

Image Segmentation	Precision [%]	Recall [%]	F1 [%]
Belief Propagation	55.34	19.47	21.03
K-Means	43.27	15.13	17.56
Hybrid segmentation	61.49	25.09	27.52

The hybrid segmentation algorithm in a combination with SIFT-SAD stereo matching method is applied to refine the final depth map (see Fig. 6).

The proposed stereo matching algorithm based on combination SIFT-SAD method using hybrid segmentation algorithm was very efficient. The resulting images were computed on the 2.27 GHz Intel Core i3 processor

with 4 GB DDR3 memory. The computational time for only SAD algorithm was approximately 125 seconds and for SIFT-SAD method 59 seconds, respectively. The stereo matching algorithm based on SIFT-SAD seems to be more effective algorithm to producing the cleaner disparity map with the homogeneous areas.



Fig. 6. Depth map of the hybrid segmentation algorithm implementation

The set of all parameters used in our hybrid segmentation algorithm is shown in Table 3. Spatial resolution parameter s affects smoothing and connectivity of segments. Moreover, parameter S is a size of the smallest segment, Min_sh is minimum and Max_sh is maximum shift of the pixels. All the experiments were practiced on pictures from 101 object images database [11].

TABLE 3. PARAMETERS USED IN HYBRID ALGORITHM

Parameter	Set value
s	5
S	50
Min_sh	1
Max_sh	40

Finally, the four methods (SIFT, ASIFT, SURF, SIFT-SAD) are compared. This all methods are based on combination OpenCV and MATLAB. We use the same dataset [11], which includes the general deformations, such as view, illumination and rotation changes. Time evaluation is a relative result, which only shows the tendency of the four methods' time cost. There are factors that influenced on the results such as the size and quality of the image, image types (e.g. scenery or texture), and the parameters of the algorithm (e.g. the distance ratio) [6].

TABLE 4. PROCESSING TIME COMPARISON

	SIFT	ASIFT	SURF	SIFT-SAD
Total matches	125	135	89	312
Total time [s]	6.82	5.07	2.78	4.95
15 matches time [s]	4.15	4.72	2.07	4.31

In this part of the experiment we uses 101 object image dataset [11], whose sizes are all 300 x 240 pixels. The parameters of the four algorithms are the same settings according to the original paper. Time is counted for the complete processing which includes feature detecting and matching. Total matching time (see Table 4) is the

computational time of finding all matches. The computational time for SIFT descriptor was approximately 7 seconds and for proposed SIFT-SAD method 5 seconds, respectively. Table 4 shown that SURF is fastest one, SIFT-SAD is slower but it detects so many key points and finds most matches. Furthermore, the proposed method is noise insensitive.

The comparison among the three algorithms, the experimental results show that the proposed SIFT-SAD method is far efficient than SIFT, ASIFT or SURF algorithm.

VI. CONCLUSION

This paper describes a procedure for recovering depth map based on the proposed hybrid segmentation algorithm using process filtering. The method for reconstructing a 3D scene from two input images was presented. We mentioned some manners allowing a three - dimensional reconstruction of picture or object in this article. The proposed system is based on 3D reconstruction solution using stereo images. This system works with common cameras. The applications of these methods of 3D picture processing are very useful in sphere of medicine, for example detection and identification of tumor in brain and also in other branches as physics, astronomy, biology or geography.

In future work, we plan to perform experiments (we could speed up computation time and improve precision of hybrid segmentation algorithm) and also tests of more complex algorithms on greater number of real images with aim to compare the presented approach with other existing algorithms.

ACKNOWLEDGEMENTS

This paper has been supported by the Slovak Science project Grant Agency, Project No. 26220220407 "Competence center for research and development of diagnostics and therapy of cancer".

REFERENCES

- [1] J.S. Yedida, W.T. Freeman, Y. Weiss (2003) Understanding belief propagation and its generalizations. *Exploring Artificial Intelligence in the New Millennium. Chap. 8. pp 239-236.*
- [2] J. Sun, N.N. Zhang, H.Y. Shum (2003) Stereo matching using belief propagation. *In IEEE Transactions on Pattern Analysis and Machine Intelligence. 25(7):787-800. 2003.*
- [3] L. Dong, P. Ogunbona, W. Li, G. Yu, L. Fan, G. Zheng (2006) A fast algorithm for color image segmentation, *First International Conference on Innovative Computing, Information and Control. 2006.*
- [4] Z. Yin, R. Collins (2007) Belief Propagation in a 3D Spatio-temporal MRF for moving object detection. Department of Computer Science and Engineering. The Pennsylvania State University. University Park PA. 2007.
- [5] M. Benco, R. Hudec (2006) The advanced image segmentation techniques for broadly useful retrieval in large image database. *NSSS Tatranske Zruby, Slovakia. ISBN 978-80-8040-344-7, pp. 40-44. 2006.*
- [6] L. Juan, O. Gwun (2009) A Comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP). Vol. 3, No. 4, pp. 143-152, 2009.*
- [7] D.G. Lowe (2004) Distinctive image feature from scale-invariant key points. *International Journal of Computer Vision (IJCV). Vol. 2, No. 60, pp. 91-110.*
- [8] A. Kuhl (2005) A Comparison of Stereo Matching Algorithms for Mobile Robots. Centre for Intelligent Information Processing System. The University of Western Australia, pp. 4-24.
- [9] M. Zachariasova, R. Hudec, M. Benco, P. Kamencay (2012) The object recognition based on scale-invariant feature transform and hybrid segmentation. *Proceedings of the 9th International Conference ELEKTRO 2012. CD ISBN 978-1-4673-1178-6.*
- [10] P. Kamencay, M. Breznan, R. Jarina, P. Lukac, M. Zachariasova (2012) Improved depth map estimation from stereo images based on hybrid method. *The Radioengineering Journal, Vol. 21, No. 1, ISSN 1210-2512.*
- [11] Computational Vision at Caltech 101 [Online] Cited 2012-07-17. Available at <http://vision.caltech.edu/>.
- [12] K. Mikolajczyk, C. Schmid (2008) Scale and Affine Invariant Interest Point Detectors. *International Journal of Computer Vision (IJCV). Vol. 60, No. 1. pp. 63-86.*
- [13] R. Hudec, M. Benco, J. Krajcovic (2008) SD LMS L-filters for filtration of gray level images in time-spatial domain based on GLCM features. *Advances in Electrical and Electronic Engineering. Zilina, Slovakia. Vol. 7, No. 1-2. pp. 51-54.*
- [14] R. Hudec (2011) Adaptive order-statistics L-filters. 1st ed. University of Zilina. EDIS Press. p. 155. ISBN 978-80-554-0248-2.
- [15] M. Shimizu, M. Okutomi (2008) Calibration and rectification for reflection stereo. *Computer Vision and Pattern Recognition CVPR. IEEE Conference. pp. 1-8.*
- [16] D. Scharstein, R. Szeliski (2002) A taxonomy and evaluation of dense two frame stereo correspondence algorithms. *International Journal of Computer Vision IJCV. pp. 7-42.*
- [17] R. Martín (2010) Towards Ground Truth Disparity Maps of Natural Scenes: Combining Laser Range Scans and Stereo Images. *Publications of the Institute of Cognitive Science. Vol. 23. 2010.*
- [18] R. Queiroz, M. Cohen, J.L. Moreira, A. Braun, J.C. Jacques, S.R. Musse (2010) Generation facial ground truth with synthetic faces. *Graphics, Patterns and Images (SIBGRAPI). pp. 25-31.*
- [19] B. McKinnon, J. Baltes (2005) Practical region-based matching for stereo vision, *IWCIA, Vol. 3322 of Lecture Notes in Computer Science, Springer, pp. 726-738.*

Evaluation of Mechanical Properties of Aluminum Alloy 6061- Glass Particulates reinforced Metal Matrix Composites

Madhu Kumar YC¹, Uma Shankar²

¹P.G student Department of Mechanical Engineering, SIT Tumkur

²Professor, Department of Mechanical Engineering, SIT Tumkur

Abstract: Aluminum matrix composites (AMCs) refer to the class of light weight high performance aluminum centric material systems. The reinforcement in AMCs could be in the form of continuous/discontinuous fibers, whisker or particulates, in volume fractions. Properties of AMCs can be tailored to the demands of different industrial applications by suitable combinations of matrix, reinforcement and processing route. This work focuses on the fabrication of aluminum alloy (6061) matrix composites (AMCs) reinforced with 3 to 12 wt% glass particulates of 75 μ m, 88 μ m, 105 μ m and 250 μ m using stir casting route. The microstructure and mechanical properties of the fabricated AMCs were analyzed. The mechanical properties like hardness and tensile strength of the unreinforced alloy and composites have been measured. The mechanical properties like hardness and tensile strength have improved with the increase in weight percentage of glass particulates in the aluminum matrix.

Keywords: Metal Matrix Composites, Glass particulate, Stir casting.

I. INTRODUCTION

Metal matrix composites (MMCs) are increasingly becoming a new class of material in aerospace applications because, their properties can be tailored through the addition of selected reinforcements [1-2]. In particular, particulate reinforced MMCs have recently found special interest because of their specific strength and specific stiffness at room and elevated temperatures [3]. Applications of Aluminum-based MMCs have increased in recent years as engineering materials. The introduction of a ceramic material into a metal matrix produces a composite material that results in an attractive combination of physical and mechanical properties which cannot be obtained with monolithic alloys. Discontinuously reinforced aluminum matrix composites have emerged from the need for light weight, high stiffness materials which are desirable in many applications, mainly on automobile products such as engine piston, cylinder liner, brake disc/drum etc. The strengthening of aluminum alloys with a reinforcement of fine ceramic particulates has greatly increased their potential in wear resistant and structural applications [1-13]. There is an increasing interest in the development of metal matrix composites (MMCs) having low density and low cost reinforcements. Although these MMCs have better properties including high strength, high stiffness and better wear resistance their usage is limited due to their high manufacturing cost. Among the various discontinuous reinforcements used, glass particulate is one of the most inexpensive and low-density reinforcement. Incorporation of glass particles reduces the cost and density of aluminum and its alloys.

A. M. S. Hamouda, S. Sulaiman, T. R. Vijayaram, M.Sayuti, M.H.M.Ahmad.[1] discusses the processing and characterization of quartz particulate reinforced aluminum-silicon alloy matrix composite which were fabricated by stir casting technique with percentages of SiO₂ particle varying from 5 to 30 wt% with particle size of 65 μ m in steps of 5 wt%. Hardness values were measured for the quartz particulate reinforced LM6 alloy composites and it has been found that it gradually increases with increased addition of the reinforcement phase. The tensile strength of the composites decreases with the increase in addition of quartz particulate.

Sudarshan, M.K. Surappa.[2] in their paper deals with the mechanical properties such as hardness, tensile strength, compressive and damping characteristics of A356 Al and A356 Al-fly ash prepared using stir-cast technique and hot extrusion, in which 6-12 wt% of fly ash was dispersed in the base matrix. Bulk hardness, matrix microhardness, 0.2% proof stress of A356 Al-fly ash composites are higher compared to that of the unreinforced alloy. Additions of fly ash lead to increase in hardness, elastic modulus and 0.2% proof stress. Composites reinforced with narrow size range fly ash particle exhibit superior mechanical properties compared to composites with wide size range particles. A356 Al-fly ash MMCs were found to exhibit improved damping capacity when compared to unreinforced alloy at ambient temperature. Damping capacity of fly ash reinforced Al-based composite increases with the increase in volume fraction of fly ash.

Joel Hemanth.[5] Have studied the mechanical and abrasive, slurry erosive wear with chilling effect of fused silica (SiO_{2p}) as reinforcement and aluminum alloy (A356) as a base material. The chills used were metallic and non-metallic chills. The fused silica particles of size 50-100 μ m were used as reinforcement varying from 3-12 wt% in steps of 3 wt% Strength, hardness and wear resistance increase up to 9 wt. % additions of dispersoid and copper chill was found to be the best because of its high volumetric heat capacity.

E. Mohammad Sharifi, F. Karimzadeh, M.H. Enayati. [3] Investigated the mechanical and tribological properties of boron carbide reinforced aluminum matrix nanocomposites were fabricated by mechanical alloying with percentages of boron carbide varying from 5 to 15 wt% in steps of 5 wt%. The sample with 15 wt % B4C had the optimum properties. This sample had a value of 164 HV which is significantly higher than 33 HV for pure Al. Also, ultimate compressive strength of the sample was measured to be 485 MPa which is much higher than that for pure Al (130 MPa). The wear resistance of the nanocomposites increased significantly by increasing the B4C content. Dominant wear mechanisms for Al-B4C nanocomposites

were determined to be formation of mechanical mixed layer on the surface of samples.

In the present investigation, aluminum based metal matrix composite containing 3 to 12 wt% of glass particulates were successfully synthesized using stir casting method. Evaluating the mechanical properties of produced composites.

II. Materials

2.1 Matrix material

The matrix material used in the experimental investigation was an aluminum alloy (6061) whose chemical composition is listed in Table 1. It therefore has a low melting point 660°C. Aluminium alloy in its unmodified state is extensively used in sand casting and die-casting. The molten metal has high fluidity and solidifies at constant temperature.

Table.1 Chemical composition of Aluminum Alloy 6061 by wt%

Cu	Mg	Si	Fe	Mn	Cr	Zn	Aluminum
0.4	1.2	0.80	0.70	0.15	0.35	0.25	balance

2.2 Reinforcement material

The reinforcement material used in the investigation was glass particulates of particle size of 75µm, 88µm, 105µm and 250µm. Particulates size was estimated by sieve analysis. Chemical composition of reinforcement is listed in Table 2. Particle density is of 2.44 g/cm³ and melting point is of 1400°C.

Table.2 Chemical composition of Glass particulates by wt%

SiO ₂	Na ₂ O	CaO	MgO	Al ₂ O ₃	K ₂ O	TiO ₂	Fe ₂ O ₃
73	14	9	4	0.15	0.03	0.02	0.1

III. Experimental Procedure

The synthesis of the metal matrix composite used in the present work was carried out by stir casting route. Al alloy was used in the form of ingots. The cleaned metal ingots were melted to the desired temperature of 740°C in graphite crucibles. Cover flux was added in to the molten metal in order to minimize the oxidation. Electrical resistance furnace with temperature controlling device was used for melting. For each melting 0.250 kg of alloy was used. 3g of C₂Cl₆ – solid hexachloro ethane was added as degassing tablet in to the super heated molten metal at a temperature of 700°C. Glass particulates preheated to around 300°C for 30 mins were then added to the molten metal and stirred continuously for 5 min. During stirring, magnesium was added in small quantities to increase the wettability of glass particulates. The dispersion of the preheated glass particulates was achieved in accordance with the stir casting route. The melt with reinforced particulates were poured into the dried cylindrical permanent metallic moulds of size 12.5mm diameter and 160mm height. The pouring temperature was maintained at 680°C. The melt was allowed to solidify in the moulds.

3.1 Tensile strength

The tensile tests were conducted on servo hydraulic UTM at room temperature. The samples were prepared according to ASTM E8M. The tensile properties of the alloys were determined by performing the tension test on standard cylindrical tensile specimens. A typical tensile specimen as per ASTM standard is shown in Fig 1.

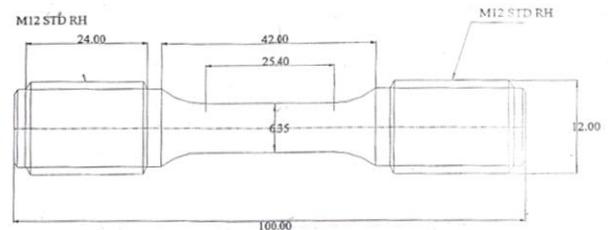


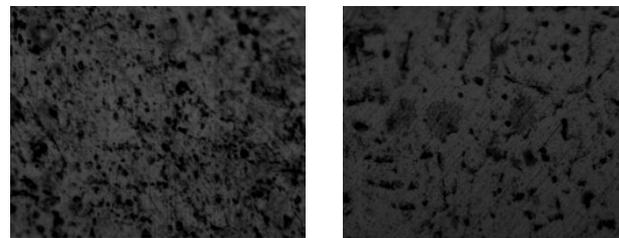
Fig.1. Tensile Specimen as per ASTM standard

3.2 Hardness testing

Micro hardness was calculated by using Zwick/Roell Micro Vicker's hardness testing machine. A precision diamond indenter is impressed on material at a load of 50 grams for 10 secs. In order to avoid the segregation effect of the particles, fifteen readings were taken for each sample and the average value is reported.

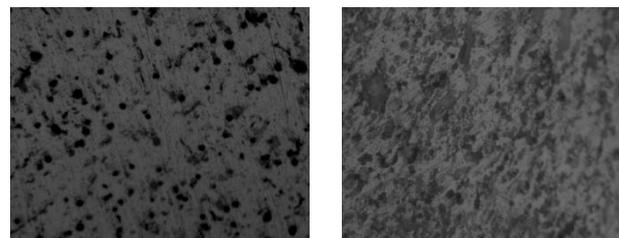
IV. Results And Discussion

4.1. Evaluation of microstructure



(a)

(b)



(c)

(d)

Fig.3. Photomicrographs of the cast Al6061-75µm glass particulate AMCs: (a) 3% glass particulate, (b) 6% glass particulate, (c) 9% glass particulate, (d) 12% glass particulate at 100x.

The optical photomicrographs of the fabricated AMCs are shown in Fig. 3. It is observed from the figure that glass particulate are dispersed uniformly in the aluminum matrix at all weight percentage. The size of the glass particles appears to be uniform throughout the aluminum matrix. This can be attributed to the effective stirring action and the use of appropriate process

parameters. Homogeneous distribution of particles is to enhance the mechanical properties of the matrix alloy.

4.2. Evaluation of tensile strength

Fig. 4 shows the relation between weight percentage of glass particulates and tensile strength of fabricated composites. It is observed that the Al alloy (6061) has tensile strength of 119 Mpa. Tensile strength increases by adding the reinforcement up to 9wt% and then decreases with increasing wt% of glass particles. Increase in tensile strength (reinforcement up to 9 wt. %) is attributed to increase in grain boundary area due to grain refinement, at the interface and effective transfer of applied tensile load to the uniformly distributed well bonded reinforcement.

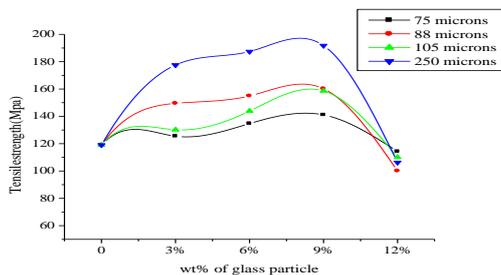


Fig.4. Shows the relation between weight percentage of glass particulates and tensile strength

4.3. Evaluation of hardness

Fig. 5. Shows the relation between weight percentage of glass particulates and hardness of fabricated composites. It is observed from the figure that the hardness of Al alloy (6061) is 75 and then hardness increases with increasing wt% of glass particulate upto 9wt% and then decreases with increasing wt% of glass particles.

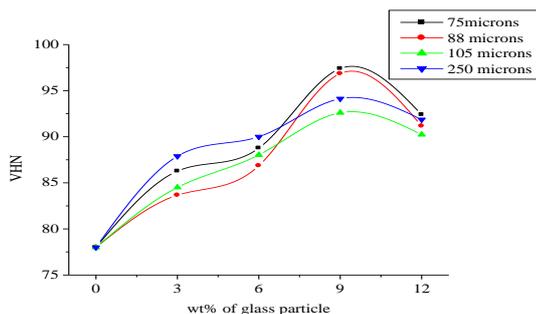


Fig.5. Shows the relation between weight percentage of glass particulates and hardness.

V. Conclusions

Aluminum-Glass particulate composite was successfully synthesized by the stir casting method. The tensile strength of the composite increased with increase in wt% of glass particulates up to 9%. The micro hardness of the composites increased with increase in wt% of the dispersoid up to 9 wt% and further addition of dispersoid showed that the hardness decreases. Microstructural observations show that the glass particulates are uniformly distributed in the Al6061 matrix.

References

- [1] A.M.S.Hamouda, S.Sulaiman, T.R Vijayaram, M.Sayuti, M.H.M.Ahmad, Processing and characterization of particulate reinforced aluminum silicon matrix composite *journal of achievements in materials and manufacturing engineering, volume 25 issue2*, December 2007.
- [2] Sudarshan, M.K. Surappa, Synthesis of fly ash particle reinforced A356 Al composites and their characterization, *Materials Science and Engineering A* 480, 2008 117–124.
- [3] E. Mohammad Sharifi, F. Karimzadeh, M.H. Enayati, Fabrication and evaluation of mechanical and tribological properties of boron carbide reinforced aluminum matrix nanocomposites, *Materials and Design* 32, 2011 3263–3271.
- [4] H.C. Anilkumar, H.S. Hebbar, K.S. Ravishankar, Mechanical Properties of Fly Ash Reinforced Aluminium Alloy (Al6061) Composites, *International Journal of Mechanical and Materials Engineering* Vol.6 , No.1, (2011) 41-45.
- [5] Joel Hemanth, Quartz (SiO₂p) reinforced chilled metal matrix composite (CMC) for automotive applications, *Materials and Design* 30, (2009) 323–329.
- [6] M.Sreenivasa Reddy, Soma V. Chetty, Sudheer Premkumar, Effect of reinforcements and heat treatment on tensile strength of Al-Si-Mg based hybrid composites, *Int. Journal of Applied Sciences and Engineering Research, Vol. 1, No. 1*, 2012.
- [7] K. Kalaiselvan, N. Murugan, Siva Parameswaran, Production and characterization of AA6061–B4C stir cast composite, *Materials and Design* 32, (2011) 4004–4009.
- [8] E. Bernardo, G. Scarinci, A. Maddalena, S. Hreglich, Development and mechanical properties of metal particulate glass matrix composites from recycled glasses, *Composites Part A*, 35, (2004) 17–22 .
- [9] J. Babu Rao, D. Venkata Rao, N.R.M.R. Bhargava, Development of light weight ALFA composites, *International Journal of Engineering, Science and Technology* Vol. 2, No. 11, 2010, pp. 50-59.
- [10] Ayhan Erol, Ahmet Yonetken, Ismail Yıldız, Muzaffer Erdogan, Fabrication of Ni Metal Matrix Composites Reinforced With SiO₂ by Microwave Sintering, *International Advanced Technologies Symposium*, May 13-15, 2009.
- [11] Manoj Singla, D.Deepak Dwivedi, Lakhvir Singh, Vikas Chawla, Development of Aluminium Based Silicon Carbide Particulate Metal Matrix Composite, *Journal of Minerals & Materials Characterization & Engineering, Vol. 8, No.6*, 2009, pp 455-467.
- [12] G. B. Veeresh Kumar, C. S. P. Rao, N. Selvaraj, Mechanical and Tribological Behavior of Particulate Reinforced Aluminum Metal Matrix Composites – a review, *Journal of Minerals & Materials Characterization & Engineering, Vol. 10, No.1*, 2011 pp.59-91.
- [13] A. R. K. Swamy, A. Ramesha, G.B. Veeresh Kumar, J. N. Prakash, Effect of Particulate Reinforcements on the Mechanical Properties of Al6061-WC and Al6061-Gr MMCs, *Journal of Minerals & Materials Characterization & Engineering, Vol. 10, No.12*, 2011, pp.1141-1152.
- [14] V. Senthilkumar, A. Balaji, Hafeez Ahamed, Effect of Secondary Processing and Nanoscale Reinforcement on the Mechanical Properties of Al-TiC Composites, *Journal of Minerals & Materials Characterization & Engineering, Vol. 10, No.14*, 2011, pp.1293-1306.
- [15] S. Sarkar, Akhilesh Singh, Studies on Aluminum-Iron Ore in-Situ Particulate Composite, *Open Journal of Composite Materials* 2, 2012, 22-30.

Low Flow Characterization of a Coastal River in Ghana

E. O. Bekoe¹, F. Y. Logah², K. Kankam-Yeboah³, B. Amisigo⁴

^{1,2}Research Scientists, ³Principal Research Scientist and ⁴Senior Research Scientist

^{1,2,3,4}Water Research Institute, Council for Scientific and Industrial Research, P. O. Box M32, Accra, Ghana.

Abstract: Various probability distribution functions including Normal, Lognormal, Weibul, Gumbel and Gamma distributions were fitted to the mean daily low streamflows for the coastal river Ayensu in Ghana to characterize the low flow regime of the river. The Normal and Gumbel distributions produced the best fit with NSE equaled to 99.17% and 99.19%, respectively. A Flow Duration Curve was developed and used to determine the minimum flow threshold for the Ayensu River using mean daily streamflow series at Okyereko gauging station. Results showed that streamflow in the basin at Okyereko had little tendency to produce unusual extreme low flow with the minimum flow threshold value of 0.20 m³/s which is equaled or exceeded 95% of the time. The probability of occurrence of low extreme flows in the basin is low and that water abstraction in terms of use for water supply for domestic, industrial and agricultural requirement is considered reliable and sustainable.

Keywords: Low flow, flow duration curve, Weibul-Gumbel distribution, Ayensu river basin, Ghana

I. Introduction

Low streamflow statistics, according to [1], indicate the probable availability of water in streams during times when conflicts between water supply and demand are most likely to arise. Because of this, low streamflow statistics are needed by the state, regional and local agencies for water-use planning, management and regulatory activities for a variety of water resources application. These activities include (i) developing environmentally sound river-basin management plans, (ii) siting and permitting new water withdrawals, inter-basin transfers and effluent discharges, (iii) determining minimum streamflow thresholds for the maintenance of aquatic biota and (iv) land-use planning and regulation. Continuous water supply demands continuous abstraction from the surface and ground water bodies. In abstracting water from rivers, consideration should be given to the minimum flow needed to sustain the stream. Also, it is important to determine the reliability of streams to water supply during the dry seasons where the amount of river flow is low.

Estimation of low streamflow statistics at gauged river sites involve evaluation of annual n-day minimum streamflow, description of annual minimum streamflow through the selection of a probability distribution and the estimation of the distribution's parameters [2]. Low flow conditions of a stream may be described by several low streamflow characteristics in the form of indices and exceedance percentile. Depending on the type of data initially available and the type of output information required, there exist different methods for estimating low-streamflow indices. These include Flow Duration Curve (FDC), Low Streamflow Frequency Analysis (LSFA) and Flow Distribution Functions (FDF).

Studies [3] conducted on water resources in Ghana showed that the country is endowed with sufficient surface water resources to serve all its water needs. However, there is the need for a gradual process of development and conservation to make the water available in sufficient quantity and good quality [3],[4]. Yet in the dry seasons some rivers dry up and hinder certain water uses such as agriculture, domestic water supply, navigation and hydropower generation. Thus low flow statistics are needed to determine the availability of water for water supply, waste discharge and power generation. According to [5], the assessment of low streamflow is important because it is a critical index for these water projects. The Ayensu River being part of the Coastal river systems of Ghana, is being characterized because of its economic importance [6]. According to [6] a baseline survey conducted in 1997 in the Ayensu basin identified inadequate water supply as one of the problems facing the irrigation scheme. Furthermore, [7] established that water delivery flexibility index for the project area was 5 and tail-end supply ratio of 0.45 was noted. Further [8] reported a high water stress/vulnerability index for the basin beyond 2020.

Thus, this paper sets out to use the probability distributive functions namely, Gumbel, Weibul, Log-normal, Gamma and normal to model the low flow regime of the river to establish the best fit to characterize the low flow regime. This will enable the properties of flow for the river to be established to compliment better management of the basin.

II. Study Area

The Ayensu river basin (Figure 1) is part of the Coastal river system of Ghana with an area of approximately 171 km² and length of 98km² [9]. It lies between latitudes 5°20'N to 6°05'N and longitude 0°30'W to 0°50'W. The main tributary of the river is Akora [10]. The basin is located in two climatic regions; i.e. the wet Semi-Equatorial in the northern.

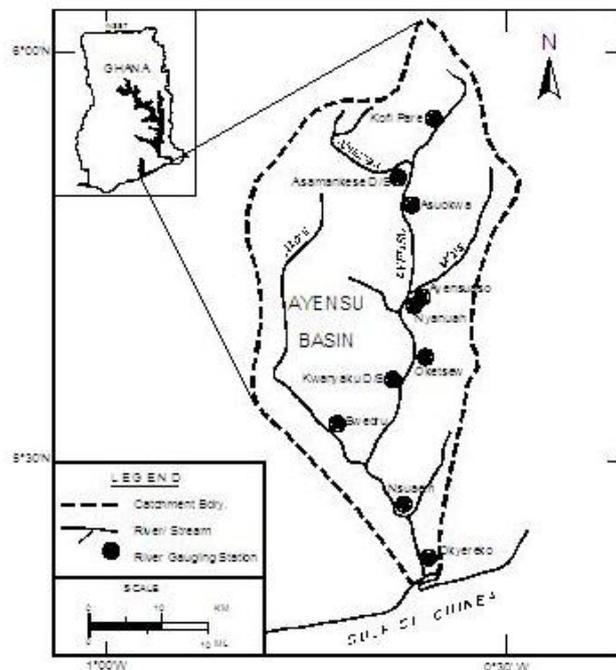


Figure 1: Map of Ayensu River Basin

Part and the dry Equatorial in the south. Rainfall in the basin is seasonal, with two rainfall peaks in June and September, where dry periods span between December and March. However, the dry Equatorial region has mean annual rainfall less than 900 mm while the wet Equatorial has a mean annual rainfall between 1200 mm and 2000 mm [10]. The Ayensu river is perennial suggesting that groundwater plays a very important role in its existence. This ground water resource in Ayensu river basin is fresh [11]. The dominant soil type is forest ochrosols, which covers about 95% of the area. The other soil type is savannah ochrosols and savannah lithosols in the southern part of the basin. Three vegetations types are found in the basin. The upper and the middle parts are covered by moist semi-deciduous forest. The remaining third of the basin is coastal thickets and grasslands. The mean annual stream flow [9] is $8.27\text{m}^3/\text{s}$ with maximum flows occurring between June-July with mean annuals of $20.89\text{-}22.40\text{m}^3/\text{s}$. Annual runoff is estimated [9] to be 268 million m^3/s .

The Ayensu river basin harbours two important water schemes. i.e. the Okyereko Irrigation scheme and the Kwanyako Water Supply System (Kwanyako Headworks Project) in the Central Region. The dam and water supply system at Kwanyako was established in 1964 to supply treated water for the surrounding communities. It was rehabilitated in 1998 and 2005 which increased the total water supply capacity of the system from $12,440\text{m}^3/\text{day}$ to $35,000\text{m}^3/\text{day}$ [12]. Currently the system serves 13 towns and 160 surrounding communities including Cape Coast in five (5 No.) districts in the Central Region at an average production rate of $90,000\text{m}^3$ of water per day. The Okyereko Irrigation Scheme was constructed between 1973-1982 and rehabilitated (1996-2004) as a pilot scheme under the Small-Scale Irrigated Agriculture Promotion Project (SSIAPP) to support local agriculture in the basin.

III. Methodology

Statistical analyses which according to [13] are widely applied to derive indices to characterize low streamflow regimes are the main tools used to characterize the coastal catchment. All analyses were done in MS Excel.

3.1 Streamflow data

The basic data used for the study was the mean daily streamflow data series collected from the Ayensu basin at the Okyereko river station in Ghana. This data set was used because of its relatively good data length and continuity compared to the other stations within the basin. Available streamflow data from the station were from 1962 – 1997. They were obtained from the Hydrological Service Department (HSD) of the Ministry of Water Resources, Works and Housing (MWRWH), Accra.

3.2 Estimation of Low Streamflow

The duration of streamflow data for the study was less than 50 years, thus the peak-over threshold method [14] was adopted to define the minimum flow requirement of the river. The threshold value below which all streamflows are minimum was estimated from the flow duration curve (FDC) at 95 % probability of exceedance. The FDC for the river using the complete data series was developed and then used to extract the low streamflows at probabilities of exceedance of 95% and above.

3.3 Flow Duration Curve

A Flow Duration Curve (FDC) defines the relationship between any given discharge value and the percentage of time that this discharge is equaled or exceeded [15], [16]. The FDC is developed by plotting all ranked streamflows against their rank, expressed as the percentage of the total number of time steps in the record [15].

Ranked numbers were assigned to each streamflow value with the largest flow ranked as 1 and the smallest n , where n is the total number of records. The probability of exceedance was computed using the relation in equation (1) [14]:

$$P = 100 \times \frac{r}{n+1} \quad (1)$$

Where P is the percentage of time a given flow is equaled or exceeded, n is the total number of records and r is the rank of the flow magnitude. The FDC was obtained by plotting ranked streamflows against their rank, expressed as the percentage of the total number of time steps in the record.

3.4 Extraction of Low Streamflow

The next step was to extract the low streamflow from the ranked (or sorted) flows. The extraction was done in the Microsoft Excel Worksheet by selecting, copying and pasting in a new column the streamflows that were equaled or exceeded 95 % of the time (i.e. from 95 % to 100 % probability of exceedance).

3.5 Estimation of Baseflow Contribution

Baseflow contribution to streamflow in the basin was estimated using equation (2) with the complete flow series [17]:

$$f_b = \frac{Q_{90}}{Q_{50}} \quad (2)$$

where f_b is the fraction of baseflow contributed to low streamflow and Q_{50} and Q_{90} are the streamflows which are equaled 50 % and 90 % of the time, respectively.

3.6 Flow Frequency (Return period) Analysis

In developing the flow frequency curve, the mean daily low river discharges for the period of record were transformed into high values by using the transformation ($X=1/x$). The transformed values were sorted in descending order of magnitude and assigned rank numbers with the largest value ranked as 1 and the lowest n , where n is the total number of record data. The recurrence interval of the streamflow with certain magnitude was computed using equation (3). The streamflow frequency curve was developed by plotting the flow discharge against the empirical return period. The return period for extremes low flow values was also computed using equation (3) [18] [19]:

$$T_c = \frac{n}{t} * \left[\frac{1}{\exp\left(-\left(\frac{x^{-1}-x_t^{-1}}{\beta}\right)\right)} \right] \quad (3)$$

On the basis of linear regressions in the exponential quantile plots, the design low streamflow for certain return period (T -years) was estimated by re-arranging equation (3) into equation (4) [18, 19], i.e.

$$x_T = x_t^{-1} + \beta(\ln(T) - \ln\left(\frac{n}{t}\right)) \quad (4)$$

where x_T is the estimated design low streamflow at T -years, x_t is the threshold value below which all streamflows are low flows, T is the return period in years, n is the period of record (in years), t is the number of extracted low streamflows and β , the calibrating parameter.

3.7 Flow Distribution Functions

The Normal, Log-normal, Weibul, Gumbel and Gamma distribution functions were used based on their common use in several literatures. The type of flow distribution for the basin was identified by calibrating and validating the distribution parameters and selecting the function that best fits the streamflow.

3.7.1 Calibration and Validation of Data Sets

In order to calibrate and validate the parameters of the distribution functions, two sets of flow data were required. Streamflow values that equalled or exceeded 90 % of the time were extracted to acquire more data for analysis in this section. The calibration and validation data sets were obtained by splitting the extracted mean daily low streamflows into two. The splitting of the data was done by first randomizing the low streamflow data so that both the calibration and validation data sets would have the same range of data sets. This was achieved in Microsoft Excel by using the **RAND()** function and following the steps below:

- (i) The extracted low flow values were entered into a new column in Microsoft Excel Worksheet
- (ii) The *rand()* function was entered in the next column
- (iii) The *rand* values were selected and sorted (either ascending or descending) by *expanding the selection*.

The randomly sorted low streamflows were then split into two data sets, calibration and validation data sets.

3.7.2 Fitting Normal Distribution to Mean Daily Low Streamflows

The function *NORMDIST*($x, \mu_x, \sigma_x, 1$) was used to estimate the probability of exceedance $F_e(x)$ of a normal distribution function using equation (5).

$$F_e(x) = 1 - \text{NORMDIST}(x, \mu_x, \sigma_x, 1) \quad (5)$$

The initial parameters of the distribution, μ_x and σ_x , were estimated from the low streamflows using equations (6) and (7), respectively.

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (6)$$

$$\sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2 \quad (7)$$

3.7.3 Fitting Lognormal Distribution to Mean Daily Low Streamflows

The *NORMDIST* ($\ln x, \mu_{\ln x}, \sigma_{\ln x}, 1$) function was used to evaluate the cumulative distribution function $F_e(x)$ of a log-normal distribution function using equation (8).

$$F_e(x) = 1 - \text{NORMDIST}(\ln x, \mu_{\ln x}, \sigma_{\ln x}, 1) \quad (8)$$

The initial parameters of the distribution, $\mu_{\ln x}$ and $\sigma_{\ln x}$, were estimated from the low streamflows using equations (9 and (10), respectively.

$$\mu_{\ln x} = \frac{1}{n} \sum_{i=1}^n \ln x_i \quad (9)$$

$$\sigma_{\ln x}^2 = \frac{1}{n-1} \sum_{i=1}^n (\ln x_i - \mu_{\ln x})^2 \quad (10)$$

3.7.4 Fitting Gamma Distribution to Mean Daily Low Streamflows

From equation (11), the *GAMMADIST*($x, \lambda, k, 1$) [14] function was used to evaluate the cumulative .distributive .function of the gamma distribution function $F_e(x)$. The initial guess distribution parameters λ and k were estimated from the mean, μ and standard deviation, σ of low streamflows using equations (12) and (13), respectively.

$$F_e(x) = 1 - \text{GAMMADIST}(x, \lambda, k, 1) \quad (11)$$

$$\mu_x = \frac{k}{\lambda} \quad (12)$$

$$\sigma_x^2 = \frac{k}{\lambda^2} \quad (13)$$

3.7.5 Fitting Weibul Distribution to Mean Daily Low Streamflows

The probability of exceedance, $F_e(x)$ of a Weibul distribution function was evaluated using equation (14) [14]:

$$F_e(x) = \exp \left[- \left(\frac{x}{\beta} \right)^\tau \right] \quad (14)$$

The initial parameters of the distribution, τ and β , were estimated from the mean, μ_x and standard deviation, σ_x of low streamflows using equations (15) and (16), respectively [14].

$$\tau = \mu_x \quad (15)$$

$$\beta = \sigma_x \quad (16)$$

3.7.6 Fitting Gumbel Distribution to Mean Daily Low Streamflows

The probability of exceedance, $F_e(x)$ of a Gumbel distribution function was estimated using equation (17). The initial parameters of the distribution, x_t and β , were estimated from the mean, μ_x and standard deviation, σ_x of low streamflows using equations (18) and (19), respectively [14].

$$F(x) = 1 - \exp \left[- \exp \left(- \frac{x - x_t}{\beta} \right) \right] \quad (17)$$

$$\mu_x = x_t + 0.577216\beta \quad (18)$$

$$\sigma_x^2 = \frac{\pi^2}{6} \beta \tag{19}$$

3.8 Plotting Formula

The Weibul-Gumbel plotting position (Eq. 20) was used because it has more statistical justification and is the commonly used in hydrological frequency studies [14].

$$P = \frac{r}{n+1} \tag{20}$$

Where P is the probability that a given streamflow is equaled or exceeded, r is the order number of rank and n is the total number of records.

Once the data series was identified ranked and the plotting positions estimated, a graph of low streamflow against probability of exceedance was plotted to graphically fit a distribution function. The various distribution functions were fitted to the extracted mean daily low streamflows from the river basin. Distribution parameters were calibrated and validated with the extracted low streamflows. These were compared with the sample data to graphically observe the distribution that produced the best fit to the low streamflows in the basin.

3.9 Parameter Estimation and Optimization Technique

The accuracy or goodness of the estimated parameters was checked through the use of two main optimization techniques. These were the Root Mean Squared Error (RMSE) and the related normalization, the Nash–Sutcliffe Efficiency (NSE) [20] which according to [21] and [22] is widely used in appraising model performance: These criteria are defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n E^2}{n}} \tag{21}$$

$$NSE = 100 * \left(1 - \frac{\sum_{i=1}^n E^2}{v.n}\right) \% \tag{22}$$

$$= 100 * \left(1 - \frac{RMSE^2}{v.n}\right) \% \tag{23}$$

where n is the number of errors, v is the sample variance and E is the difference between the Weibul plotting position and the calibrated plotting positions of the distribution functions [14]. During calibration, the parameters were optimized for values which minimize the RMSE or maximize the NSE. This was achieved by using the *solver* tool in Microsoft Excel.

IV. Results and Discussions

4.1 Streamflow data

The streamflow data collected from the Ayensu basin at Okyereko is plotted (Figure2) and from this low flows were extracted. Two peak flows are usually observed (Figure 3) in the basin annually and are separated by periods of low flows with long duration. This could be as a result of the bi-modal nature of rainfall in the southern sector of the country where the Ayensu river is located.

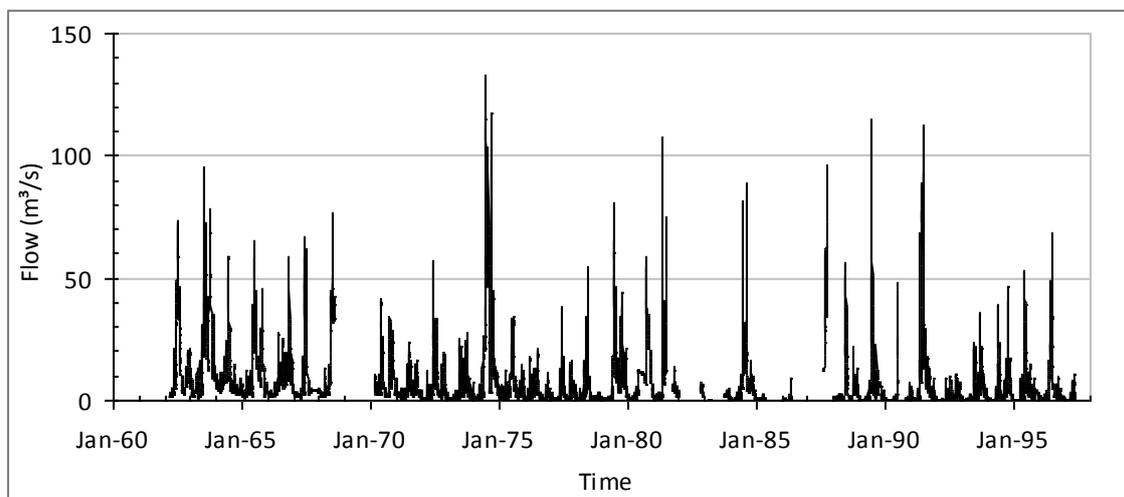


Figure 2: Streamflow series at Okyereko (1962 – 1997)

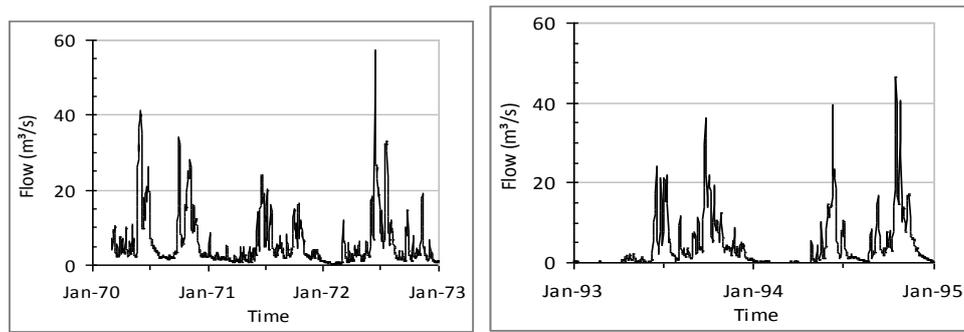


Figure 3: Fractions of historical flows at Okyereko showing the bi-modal nature of peak flows as a result of the effect of the bi-modal nature of rainfall in the southern sector of Ghana

4.2 The Flow Duration Curve and Minimum Streamflow Requirement

The mean daily low streamflows threshold value for the period of record at 95 % probability of exceedance corresponded to 0.20 m³/s (Figure 4) from the FDC and this corresponded with [9] results for the basin.

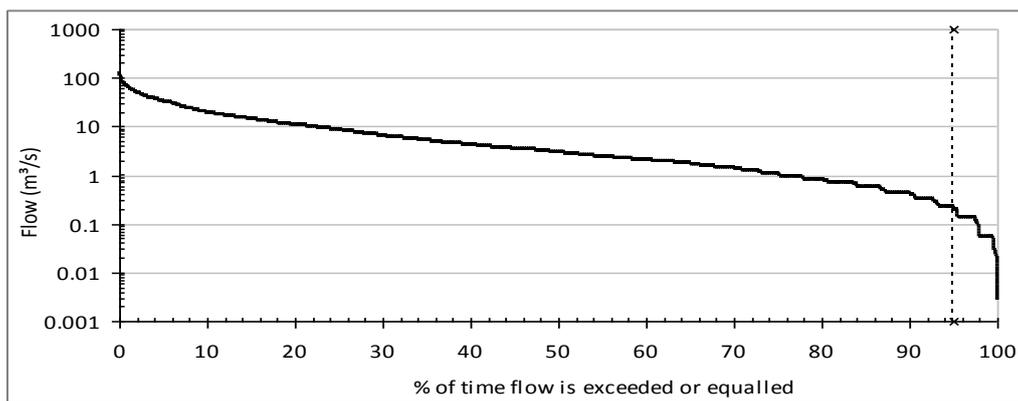


Figure 4: Flow duration curve developed for the Ayensu Basin at Okyereko using mean daily flow series.

4.3 Baseflow Index and Zero Flows

From analysis, (section 3.8) the estimated baseflow index for the period of observation was approximately 0.14 at Okyereko. This index indicated that groundwater contributed approximately 14 % to streamflow in the basin at Okyereko. This value suggested that storage of groundwater within the basin was very low. This might be due to the storage material in the basin having low permeability.

4.4 Flow Frequency (Return Period) Curve and Recurrence Intervals

Figure 5 shows the calibrated and extrapolated return period plot at Okyereko based on the exponential Extreme Value Distribution (EVD). The calibrated parameters for the river are tabulated in Table 1. From the return period plot, streamflow value of 0.100 m³/s is estimated to occur at least once every year in the basin at Okyereko. Similarly, low streamflows with magnitudes 0.016 m³/s, 0.010 m³/s and 0.009 m³/s are expected to occur at least once in a 10-year, 50-year and 100-year period, respectively.

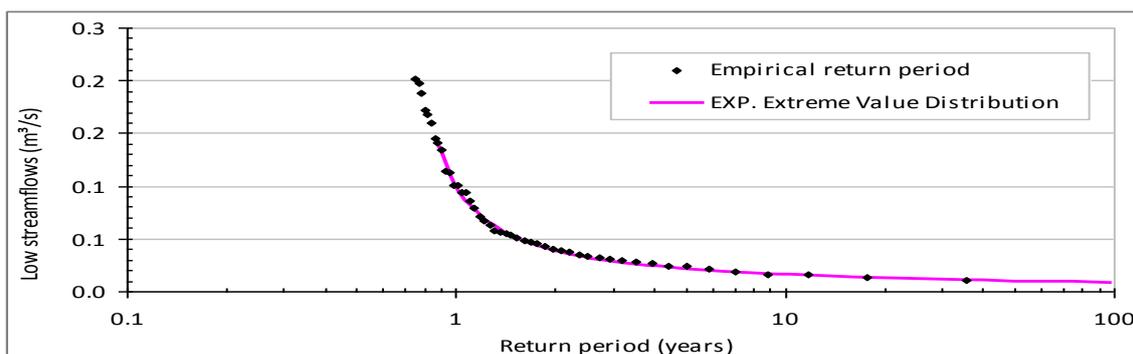


Figure 5: Return period plots for the Ayensu River Basin using low streamflows at Okyereko

Table 1: Parameter Estimates

Parameters	
Number of years of data (n)	20
Number of extracted low flows (t)	156
Threshold of low streamflow (x_t), m ³ /s	0.751
Calibrating parameter (β)	1.94

4.5 Reliability of the Okyereko River to meet future demand and supply

In Figure 6, the comparative plot between the mean monthly river flow pattern, low flow threshold line at 95 % probability of exceedance and the current water production line at Okyereko is shown. The minimum flow in the basin occurred between December and April and that the lowest flow value of 82,980 m³/day was equaled or exceeded 77.4% of the time. This value is 380 % and 860 % more than the low flow threshold value of 17,280 m³/day and the 1-year return period flow value of 8,640 m³/day, respectively, at Okyereko. However, in the month of February, the current daily water production rate (90,000 m³/day) in the basin exceeded the mean monthly flow in the basin at Okyereko by 8.5%. The flow in the Ayensu basin at Okyereko can therefore be considered sustainable and reliable in terms of use for water supply for domestic, industrial and agricultural use for the period of ten (10) months, starting from March to December (Figure 6).

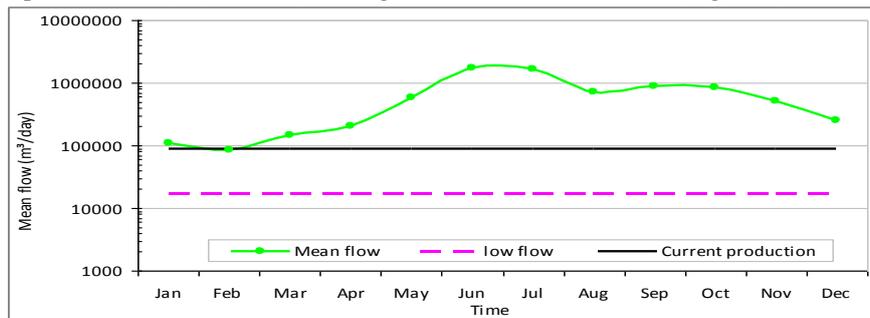


Figure 6: Graph showing the daily mean river flows in the basin, low flow line at 95% probability of exceedance and current water production line at Okyereko

4.6 Fitting flow distribution functions

Figure 7 shows the plots of the calibration and validation of low streamflow data sets for the river station. Mean daily flow data from different stations within the basin were not available for reasonable comparison to be made on which of the distributions best fitted the low streamflows in the river basin. Hence, the discussion and conclusion were based on the results obtained from streamflow data series from the Okyereko station only. Table 2 shows the values of the initial estimate and the final optimized distribution parameters for the respective distribution functions.

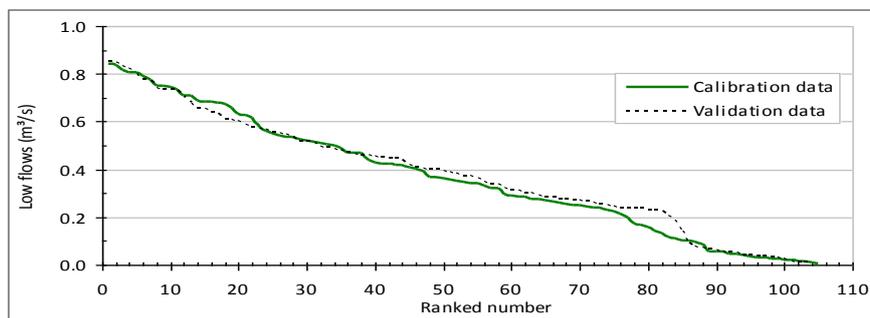


Figure 7: Calibration and Validation low streamflow data for the Ayensu Basin at Ok yereko

Table 2: Optimization of calibrated parameters for the distribution functions

Distribution functions	Parameters	Initial estimate	Optimized estimate
Normal	μ_x , (m ³ /s)	0.37	0.36
	σ_x , (m ³ /s)	0.24	0.27
Log-normal	μ_{lnx} , (m ³ /s)	-1.42	-1.12
	σ_{lnx} , (m ³ /s)	1.18	0.79
Weibul	β , (m ³ /s)	0.37	0.44
	τ , (m ³ /s)	0.24	1.40
Gumbel	β , (m ³ /s)	0.04	0.24
	x_t , (m ³ /s)	0.35	0.26
Gamma	λ , (s/m ³)	-0.14	1.68
	K	0.20	0.24

Plots of the calibrated distribution functions fitted to the mean daily low streamflows from the basin at Okyereko (Figure 8). Graphically the distribution functions fitted well with the low streamflows except for the extreme ends which was over-estimated. However, with NSE of 99.17 % and RMSE of 0.0265 m³/s the Normal distribution best fitted the mean daily low streamflows in the Basin at Okyereko. This was followed by Gumbel, Weibul, Gamma and lognormal distributions in that order.

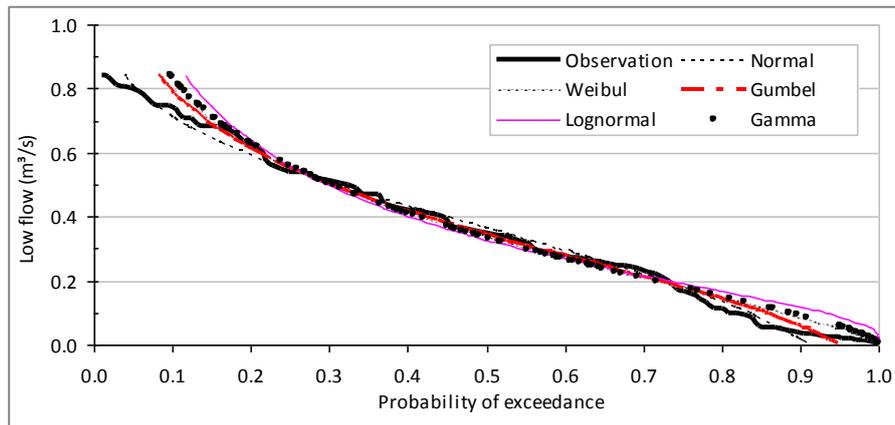


Figure 8: Calibration of distribution parameters using daily low flows from the Ayensu basin at Okyereko

It is also observed that the distribution functions fitted well with the observed mean daily low streamflows under validation mode (Figure 9) with the Normal distribution performing best with NSE of 98.87 % and RMSE of 0.305 m³/s.

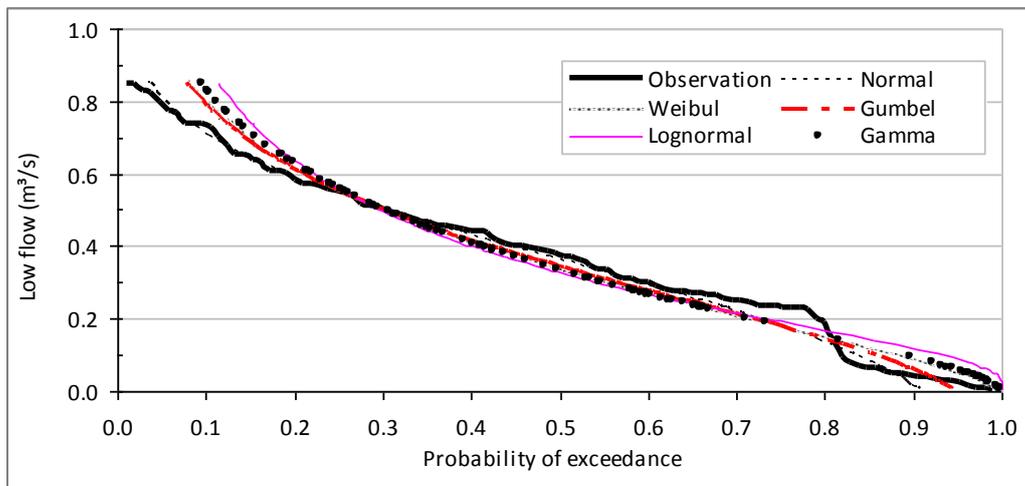


Figure 9: Validation of distribution parameters using daily low flows at Okyereko

Generally, the distribution functions fitted very well with the mean daily low streamflows but showed deviations at the extreme ends of the distributions. Apart from the Normal distribution, all the distribution functions under calibration and validation modes produced higher estimates at the extreme (lower and upper) ends of the mean daily low streamflows. This might have given an upper hand to the Normal distribution in the analysis, hence, the highest NSE and the lowest RMSE values as tabulated in Table 3.

Table 3: Statistical analysis using NSE and RMSE

Distribution functions	Calibration		Validation	
	NSE (%)	RMSE (m ³ /s)	NSE (%)	RMSE (m ³ /s)
Normal	99.17	0.0265	98.87	0.0305
Log-normal	96.17	0.0568	94.78	0.0657
Weibul	98.43	0.0363	96.98	0.0499
Gumbel	99.19	0.0261	97.97	0.0409
Gamma	97.95	0.0415	96.22	0.0559

V. Conclusion

The determination and establishment of minimum flow of streams is not only important to water users, but also very crucial for planning water supplies, managing water quality, assessing the impact of prolonged droughts on aquatic ecosystems, among others. Low flow study is essential since it educates stream users on the desirable minimum flow needed to sustain in stream uses.

Streamflow values of 0.20 m³/s was estimated from the FDC at 95 % probability of exceedance as the minimum sustainable streamflow (low flow threshold) for the flows at Okyereko in the Ayensu basin in the coastal river system of Ghana. In the most extreme case, a streamflow amount of 0.06 m³/s was equaled or exceeded 99 % of the time at Okyereko.

Groundwater contribution to streamflows in the basin was very low with an estimated baseflow index of 0.14 at Okyereko. This may be attributed to storage materials (soil and aquifer) in the basin having very low permeability. The study showed that streamflow amount of 0.100 m³/s would occur at least once every year at Okyereko in the Ayensu basin. Similarly, low streamflows with magnitudes 0.016 m³/s, 0.010 m³/s and 0.009 m³/s are expected to occur at least once in a 10, 50 and 100-year periods, respectively.

Generally, all the distribution functions under calibration and validation modes fitted very well with the mean daily low streamflows in the basin. However, the Normal and Gumbel distributions produced the best fits with NSE equaled to 99.17% & 99.19%, respectively, at Okyereko.

Low streamflow in the Ayensu basin could be described as Normal or Gumbel distributed and thus had less of a tendency to produce unusually extreme low flow at Okyereko. The probability of occurrence of low extreme flows in the basin is very low. Water abstraction from the basin below 0.20 m³/s at Okyereko is considered reliable and sustainable in terms of use for water supply for domestic, industrial and agricultural use. However with the water stress/vulnerability index for the basin beyond 2020 estimated to be high there is the need to manage this basin sustainably.

Hydrological assessment is streamflow data dependent and predictions for the future are based on historical data or information. It is therefore essential that adequate resources are set aside for the establishment of reliable monitoring stations to collect both meteorological and hydrological data to enhance scientific research in streamflow studies in the river basins of Ghana. Thus, promoting sustainable water supply for drinking, irrigation, aquaculture and fisheries, mining and manufacturing industries, ecological balance and socio-economic development of the country.

Acknowledgement

We wish to express our deepest gratitude to the Hydrological Service Department (HSD) of the Ministry of Water Resources, Works and Housing (MWRWH) of Ghana for the observed streamflow data used in this study. We are grateful to Miss Deborah Ofori of the Surface Water Division, CSIR-WRI for her useful suggestions. We would also like to thank the Technical Officers of the Surface Water Division, CSIR-Water Research Institute, for their assistance.

References

- [1] K. G. Ries III and P.J Friesz. Methods for Estimating Low Flow statistics for Massachusetts Streams 2000 <http://pubs.usgs.gov/wri/wri004135/pdf/report.pdf> Accessed 26/3/2012
- [2] C. Kroll, and R. Vogel, Probability Distribution of Low Streamflow Series in the United States. *J. Hydrol. Eng.*, 7(2), 2002, 137–146.
- [3] K. Kankam-Yeboah, P. Gyau-Boakye, M. Nishigaki and M. Komatsu, Water Resources and Environmental Management in Ghana, *Journal of the Faculty of Environmental Science and Technology, Okayama University, Vol.1.9, No.1., 2004, pp.87-98*
- [4] P.Gyau-Boakye and J. W. Tumbulto, Comparison of rainfall and runoff in the humid south-western and the semiarid northern savannah zone in Ghana. *African Journal of Science and Technology (AJST), Science and Engineering Series Vol. 7, (1), CSIR 2006, 64 – 72.*
- [5] R. H. Passchier. Low flow hydrology. Report Q3427, WL Delft Hydraulics, Delft, The Netherlands. 2004
- [6] B. D. Ofori, (Unpublished) Strategies for community participation in dam development. Issue paper submitted to International Water Management Institute (IWMI), West African Regional Office, Accra Unpublished 2008.
- [7] Y. Abukari Moses, *Irrigation performance assessment: A case study from the Okyereko Irrigation scheme, Ghana* MSc. Thesis Wageningen University, Irrigation and Water Engineering Group June MSc 2006,
- [8] WRI Water Research Institute-CSIR. *Climate Change Vulnerability and Adaptation Assessment on Water Resources of Ghana*. A UNFCC/EPA/WRI Accra report. 2000
- [9] WARM, Water Resources Management Study; Information in the Coastal basin system (Final Report). Ministry of Works and Housing, Accra, 1998
- [10] E. O. Bekoe, F. Y. Logah, B. A. Amisigo and K. Kankam-Yeboah *Climate variability and change and its impacts on water resources in Ghana*, Unpublished. Collaborative Project, UNESCO Cluster Office, Accra, Ghana, pg 37. 2010.,
- [11] N. Zakaria, T. Tetteh. Akiti, T. T., Osae, S. Adomako Dickson, S. Y. Ganyaglo, J. E. K. Hanson, and G. Ayanu, Hydrogeochemistry of groundwater in parts of the Ayensu Basin of Ghana. *Proc of the International Academy of Ecology and Environmental Sciences*, 2012, 2(2):128-135. URL: 2012. [http://www.iaees.org/publications/journals/piaees/articles/2012-2\(2\)/hydrogeochemistry-of-groundwater-in-parts.pdf](http://www.iaees.org/publications/journals/piaees/articles/2012-2(2)/hydrogeochemistry-of-groundwater-in-parts.pdf)
- [13] A. Longobardi and P. Villani 2008. 'A Regional Analysis of River's Low Flow Regime'. Department of Civil Engineering, University of Salerno, Italy, URL: www.cugri.unisa.it.
- [14] P. Willems, Statistics for Water Engineering, Chapter 3 on 'lecture notes on statistical analysis of extreme values, Katholieke Universiteit Leuven (KULeuven), Leuven, Belgium. 2000
- [15] V. U. Smakhtin. Low flow hydrology: A review, *Journal of hydrology*, 240 2001 .pp. 147-186. URL: <http://www.sciencedirect.com/science>.
- [16] L. W. Mays, Water Resources Engineering, Civil and Environmental Engineering Department, Arizona State University, Tempe, Arizona. 2005

- [17] R. J. Nathan and T. A. McMahan. Evaluation of automated techniques for baseflow and recession analysis CSIR-Water Resources Research Institute, 26(26):1990, pp. 1465-1473.
- [18] M. Mirghani, P. Willems, and J. Kabubi J. QDF relationships for low flow return period prediction, International Conference of UNESCO Flanders FIT FRIEND/Nile Project – “Towards a better cooperation”, Sharm-El-Sheikh, Egypt, 12-14 Nov. 2005, CD-ROM Proceedings, 10 p.
- [19] P. Willems, Hydrological applications of extreme value analysis, In: Hydrology in a changing environment, H. Wheater and C. Kirby (ed.), John Wiley & Sons, Chichester, vol. III, 15-25; (ISBN 0-471-98680-6). (1998)
- [20] J. E. Nash and J.E. Sutcliffe, River flow forecasting through conceptual models-Part 1: a discussion of principles. *Journal of Hydrology* 10(3), 1970, 282-290
- [21] V. J. Gupta, S. Sorooshian & P. O. Yapo, Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*. 1999.135-143
- [22] H. Niel, J. Paturol & E. Servat, Study of parameter stability of a lumped hydrologic model in a context of climatic variability. *Journal of Hydrology*, 278, 2003 213-230.

A DER Based Single-Phase Asymmetrical 27 Level Inverter Topology

V. Gopi Latha¹, K. Ratna Raju²

(¹ PG Scholar, Department of EEE, NCET, JNTUK, Jangareddy Gudem, AP, INDIA

(² HOD, Department of EEE, NCET, JNTUK, Jangareddy Gudem, AP, INDIA

ABSTRACT : This paper deals single phase 27-level multilevel inverter topology for DER's-based DC/AC conversion system. This work mainly focuses on cascaded MLI using three unequal dc sources in order to produce twenty seven-level output. In this study, a high step-up converter is introduced as a front-end stage to improve the conversion efficiency of conventional boost converters and to stabilize the output dc voltage of DERs such as photovoltaic modules for use with the simplified MLI. A conventional 27-level cascade multilevel inverter requires a combination of 13 H-bridge (single-phase full-bridge) inverter modules but the present topology, 27 level multilevel inverter is obtained by using only 3 H-bridge inverter modules with different dc voltage sources. This MLI offers strong advances such as improved output waveforms, smaller filter size, low THD, reduced volume and cost, and lower electromagnetic interference. The need of several sources on the DC side of the converter makes multilevel technology attractive for many photovoltaic applications. This paper provides an overview of DER based 27-level multilevel inverter topology and investigates their suitability for single-phase. A simulation model based on MATLAB/SIMULINK (version 7.12) is developed.

Keywords: Cascaded H Bridge inverter, High step-up converter, PV Array, THD.

I. INTRODUCTION

Conventional energy systems and high-energy factories located in the geographical locations suitable for the production of the majority of power, which is then transported to the consumption of large long-distance transmission lines to the center. System control centers monitor and control the system continuously to ensure the quality of the power. For delivering premium electric power in terms of high efficiency, reliability, and power quality, integrating interface converters of DERs such as photovoltaic (PV), wind power, micro turbines, and fuel cells into the micro grid system has become a critical issue in recent years [1]-[4]. In such systems, most DERs usually supply a dc voltage that varies in a wide range according to various load conditions. In light of public concern about global warming and climate change, much effort has been focused on the development of environmentally friendly distributed energy resources (DERs). Thus, a dc/ac power processing interface is required and is compliant with residential, industrial, and utility grid standards. Multilevel inverters can be divided into three presentable topologies; diode-clamped, flying-capacitor, and cascaded H-bridge cell [5]-[9]. Among them, cascaded H-bridge multilevel inverters have been receiving a great attention because of their merits such as minimum number of components,

reliability, and modularity. In the viewpoint of obtaining a sinusoidal output voltage multilevel inverters may increase the number of output voltage levels. However, it will need more components resulting in complexity and cost increase. To minimize these drawbacks, multilevel inverters employing cascaded transformers have been studied. Owing to the trinary characteristic of output voltage, they can synthesize high quality output voltage near to sinusoidal waves. By using a cascaded transformer, they obtain galvanic isolation between source and loads. However, the transformer may decrease the power conversion efficiency, and volume and cost will be increased. To alleviate these problems, propose a cascaded H-bridge multilevel inverter using trinary dc input source without transformers [10]. In the case of photovoltaic (PV) systems, the grid connection of many small (less than 50 kW) specifically designed inverters can be anticipated. These generators will inevitably have some impact on the voltage waveform at the point of common coupling, and thus on the network power quality. When power-electronic sources are operated in parallel, two particular effects are apparent which affect harmonic generation: attenuation and cancellation (see [18] and [19]). Attenuation occurs since the generated currents cause voltage variations that in turn affect the other sources; the impact is such as to reduce the currents causing the disturbance. Cancellation is the result of the harmonic current components of the different sources being to some extent out of phase, resulting of course in a reduction in that particular harmonic for the aggregate. The specific case of cancellation of the switching harmonics for multiple inverters, which can generally be taken to be of statistically independent phase, has been dealt with in recently published work, [20]. The objective of this paper is to study a newly constructed transformer less 27-level Multistring inverter topology for DERs. In this work, proposed inverter is reduced to multilevel inverter topology that requires only twelve active switches instead of the fifty-two required in the conventional cascaded H-bridge (CCHB) multilevel inverter [21]. In addition, among them, two active switches are operated at the line frequency. The input to the proposed prototype is obtained from PV modules. In order to improve the conversion efficiency of conventional boost converters, a high step-up converter is also introduced as a front-end stage to stabilize the output dc voltage of each DER (PV) modules for use with the simplified multilevel.

II. PV ARRAY

PHOTOVOLTAIC (PV) power supplied to the utility grid is gaining more and more visibility, while the world's power demand is increasing [11]. Not many PV systems have so far been placed into the grid due to the relatively high cost, compared with more traditional energy sources such as oil, gas, coal, nuclear, hydro, and wind. Solid-state

inverters have been shown to be the enabling technology for putting PV systems into the grid. Photons of light with energy higher than the band-gap energy of PV material can make electrons in the material break free from atoms that hold them and create hole- electron pairs. These electrons however, will soon fall back into holes causing charge carriers to disappear. If a nearby electric field is provided, those in the conduction band can be continuously swept away from holes toward a metallic contact where they will emerge as an electric current. The electric field with in the semiconductor itself at the junction between two regions of crystals of different type, called a p-n junction. The PV cell has electrical contacts on its top and bottom to capture the electrons. When the PV cell delivers power to the load, the electrons flow out of the n-side into the connecting wire, through the load, and back to the p-side where they recombine with holes [8]. Note that conventional current flows in the opposite direction from electrons.

2.1 Demands Defined by the Photovoltaic Cell(s)

A model of a PV cell is sketched in Figure. 1(a) and its electrical characteristic is illustrated in Figure.1 (b). The most common PV technologies nowadays are the mono crystalline and the multi crystalline-silicon modules, which are based on traditional, and expensive, microelectronic manufacturing processes [11]. The MPP voltage range for these PV modules is normally defined in the range from 23 to 38 V at a power generation of approximate 160 W, and their open-circuit voltage is below 45 V. However, new technologies like thin-layer silicon, amorphous-silicon, and photo Electro Chemical (PEC) are in development [11], [12]. These types of PV modules can be made arbitrarily large by an inexpensive “roll-on-roll-off” process.

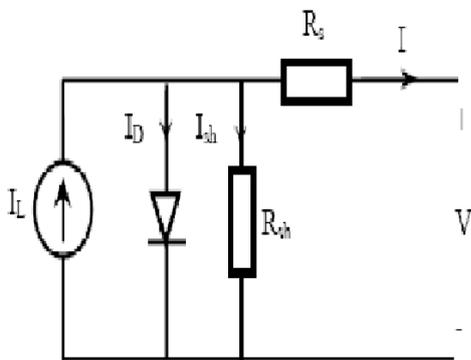


Fig 1

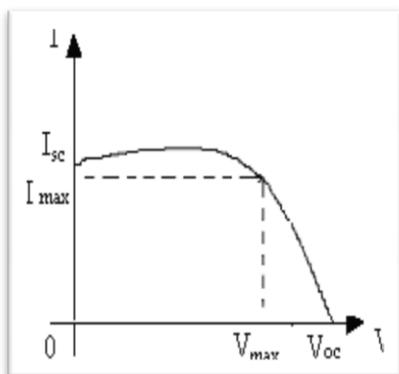


Fig 1 (b)

Figure. 1. Model and characteristics of a PV cell. (a) Electrical model with current and voltages defined. (b) Electrical characteristic of the PV cell, exposed to a given amount of light at a given temperature.

As indicated, ripple at the PV module’s terminals results in a somewhat lower power generation, compared with the case where no ripple is present at the terminals. This means that new modules with only one cell may see the light in the future. The voltage range for these cells/modules is located around 0.5, 1.0 V at several hundred amperes per square meter cell [13]–[15]. The inverters must guarantee that the PV module(s) is operated at the MPP, which is the operating condition where the most energy is captured. This is accomplished with an MPP tracker (MPPT). It also involves the ripple at the terminals of the PV module(s) being sufficiently small, in order to operate around the MPP without too much fluctuation. Analyses of the circuit in Figure. 1(a) shows that there is a relationship between the amplitude of the voltage ripple and the utilization ratio [16]. A solar cell basically is a p-n semiconductor junction. When exposed to light, a current proportional to solar irradiance is generated. The circuit model of PV cell is illustrated in Figure.1 (a). Standard simulation tools utilize the approximate diode equivalent circuit shown in Figure.1 (a) in order to simulate all electric circuits that contain diodes.

2.2 Theoretical Mathematical Model

The equation [1] & [2] that are used to solve the mathematical model of the solar cell based on simple equivalent circuit shown in Figure. 1, are given below;

$$I_D = I_o [e^{\frac{q(V+I R_s)}{K T}} - 1] \dots\dots\dots (1)$$

$$I = I_L - I_o [e^{\frac{q(V+I R_s)}{K T}} - 1] - \frac{V+I R_s}{R_{sh}} \dots\dots\dots (2)$$

Where:

- I* is the cell current in (A).
- q* is the charge of electron = 1.6x10⁻¹⁹ (coul).
- K* is the Boltzmann constant (j/K).
- T* is the cell temperature (K).
- IL* is the light generated current (A).
- I_o* is the diode saturation current.
- Rs*, *Rsh* are cell series and shunt resistance (ohms).
- V* is the cell output voltage (V).

Equation (1) was used in computer simulation to obtain the output characteristics of a solar cell, as shown in the Figure1(b). This curve clearly shows that the output characteristics of a solar cell are non linear and are crucially influenced by solar radiation, temperature and load condition.

2.3. Variation in Available Energy Due To Sun’s Incident Angle:

PV cell output with respect to sun’s angle of incidence is approximated by a cosines function at sun angles from 0° to 50° .Beyond the incident angle of 50° the available solar energy falls of rapidly as shown in the Figure 1. Therefore it is convenient and sufficient within the normal operating range to model the fluctuations in photocurrent (*I_{ph}*) verses incident angle is given by Equation (3)

2.4. Power Vs Voltage Characteristics:

Figure 2 shows the typical Power versus Voltage curve of the PV array. In this Figure, P is the power extracted from PV array and V is the voltage across the terminals of the PV array. The characteristics have different slopes at various points. When maximum power is extracted from PV array the system is operating at MPP where slope is zero. The PV curve varies according to the current insulation and temperature. When insulation increases, the power available from PV array increases whereas when temperature increases, the power available from PV Array decreases.

$$I_{ph} = I_{max} \cos \theta \quad \dots\dots\dots (3)$$

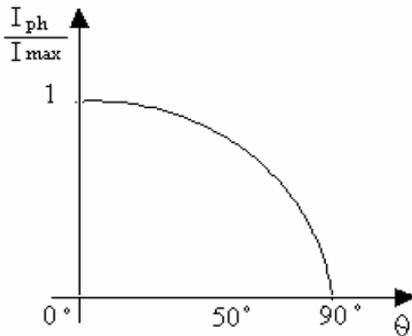


Figure 2: Power Vs Voltage

The graph shown in Figure.3 is used to find the maximum power extracted from the sun when the PV arrays are inclined a different angles. From the Figure we observe that Max power is obtained when the slope of the PV array is equal to zero.

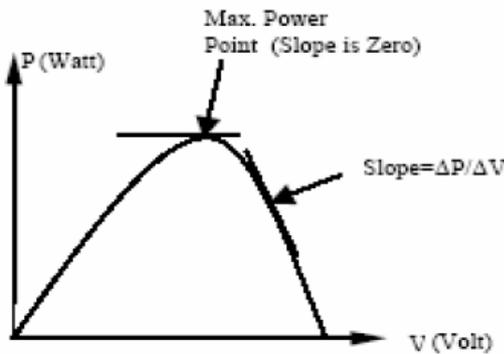


Figure. 3: Variation In Available Energy Due Sun's Incident Angle Variation.

III. HIGH STEP-UP CONVERTER STAGE

In this study, high step-up converter topology in [13] is introduced to boost and stabilize the output dc voltage of various DERs such as PV cell modules for employment of the proposed simplified multilevel inverter. The architecture of a high step-up converter initially introduced from depicted in Figure-4, and is composed of different converter topologies: boost, fly back, and a charge circuit.

The coupled inductor of the high step-up converter can be modeled as an ideal transformer, a magnetizing inductor and leaky inductor.

According to the voltage–seconds balance condition of the magnetizing inductor, the voltage of the primary winding can be derived as

$$V_{pri} = V_{in} \cdot \frac{D}{(1-D)} \quad \dots\dots\dots (4)$$

where V_{in} represents each the low-voltage DC energy input sources, and voltage of the secondary winding is

$$V_{sec} = \frac{N_s}{N_p} \cdot V_{pri} = \frac{N_s}{N_p} \cdot V_{in} \cdot \frac{D}{(1-D)} \quad \dots\dots(5)$$

Similar to that of the Boost converter, the voltage of the charge-pump capacitor C pump and clamp capacitor Cc can be expressed as

$$V_{cp} = V_{cc} = V_{in} \cdot \frac{D}{(1-D)} \quad \dots\dots\dots (6)$$

Hence, the voltage conversion ratio of the high step-up converter, named input voltage to bus voltage ratio, can be derived as

$$\frac{V_{si}}{V_{in}} = (2 + \frac{N_s}{N_p} \cdot \frac{D}{(1-D)}) \quad \dots\dots\dots (7)$$

IV. NEW MULTILEVEL INVERTER

The new hybrid multilevel inverter consists of full bridge modules which have the relationship of 1v, 3v, 9v,.....3s-1V for dc link Voltage The output waveform has 27 levels, $\pm 13 V_{dc}, \pm 12 V_{dc}, \pm 11 V_{dc}, \pm 10 V_{dc}, \pm 9 V_{dc}, \pm 8 V_{dc}, \pm 7 V_{dc}, \pm 6 V_{dc}, \pm 5 V_{dc}, \pm 4 V_{dc}, \pm 3 V_{dc}, \pm 2 V_{dc}, \pm 1 V_{dc}$, and 0. The inverter generates 3s different voltage levels (e.g. an inverter with $S = 3$ cells can generate $3^3=27$ different voltage level). The basic hybrid multilevel inverter structure for single phase is illustrated in Figure 4. This multilevel inverter is made up of a set of series connected cells. Each cell consists of a 4-switch H-bridge voltage source inverter. The output inverter voltage is obtained by summing the cell contributions. In conventional method, low level inverter is used. Better sinusoidal output was not obtained which is the drawback of the conventional system and the harmonics was high. So increase the levels of the inverter to get high resolution, hence the output wave form is mostly sinusoidal wave form. The common function of multilevel inverter is to synthesize a desired voltage from several separate DC sources [2]. Each source is connected to a single phase full bridge inverter. Each inverter is capable of generating three different output voltages, $+V_{dc}, 0$ and $-V_{dc}$.

4.1 Modeling Of New Multilevel Inverter:

For each full bridge inverter the output voltage is given by $V_{0i} = V_{dc} (S_{1i} - S_{2i})$

And the input dc current is $I_{dci} = I_a (S_{1i} - S_{2i})$
 $i = 1, 2, 3 \dots$ (Number of full bridge inverters employed). I_a is the output current of the new inverter. S_{1i} and S_{2i} is the upper switch of each full bridge inverter. Now the output voltage of each phase of the new multilevel inverter is given by

$$V_{on} = \sum_{i=1}^n V_{0i}$$

V. NEW THREE H-BRIDGE 27-LEVEL MULTILEVEL INVERTER

The topology of the proposed Dc–Ac H-bridge multilevel inverter is shown in Figure.4. The inverter uses a standard three-leg and an H bridge with its dc source in series with each phase leg.

In the proposed method of the inverter, there are three input stages. All the modules are connected as new hybrid with each module having power switches. The power

switches may be IGBT, MOSFET or any other power devices. The MOSFET's are used in this system. The power switches are operated in switching mode such that any two switches are in operating conditions at a time and other two remain in open condition. The switching is as $S_1=S_3$ and $S_2=S_4$. This method is adopted to protect the circuit from short circuiting. The number of levels is increased by connecting maximum number of modules.

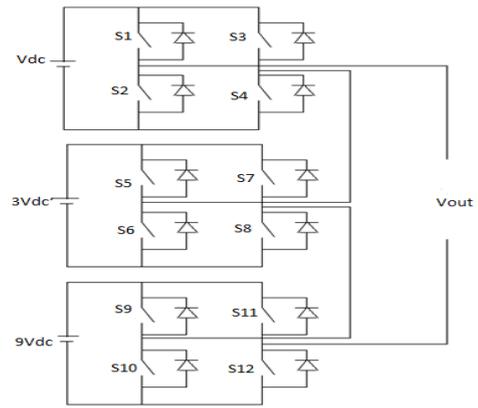


Figure 4: Single phase dc-ac three H-bridge 27 levels multi level inverter.

5.1 New Multilevel Inverter Switching:

Table.1 New Multi Inverter Switching Sequence

Output voltages and switching states for the new hybrid inverter, S=3																											
V _{dc}	V _{out}	-	-	-	-	-	-	-	-	-	-	-	0	1	2	3	4	5	6	7	8	9	10	11	12	13	
dc	13v	1v	1v	1v	9v	9v	7v	6v	5v	4v	3v	2v	1v	0v	1v	2v	3v	4v	5v	6v	7v	8v	9v	10v	11v	12v	13v
1v _{dc1}	N	O	P	N	O	P	N	O	P	N	O	P	N	O	P	N	O	P	N	O	P	N	O	P	N	O	P
3v _{dc1}	N	N	N	O	O	O	P	P	P	N	N	N	O	O	O	P	P	P	N	N	N	O	O	O	P	P	P
9v _{dc1}	N	N	N	N	N	N	N	N	N	O	O	O	O	O	O	O	O	O	P	P	P	P	P	P	P	P	P

VI. PROPOSED CONCEPT

The proposed prototype consists of three stages of operation as shown in figure 5. First stage deals with DERs such as PV modules supply dc voltage that varies in a wide range according to various load conditions. In this paper three individual dc voltages 30V, 90V, 270V are obtained from three individual PV modules and then they are boosted to 100V, 300V, 900V respectively by using boost converter in the second stage of operation. In the third stage, boosted dc sources are applied as inputs to the proposed asymmetrical 27-level multilevel topology. The proposed inverter topology requires only 12 active switches instead of the 52 required in the conventional cascaded H-bridge [21]. In order to improve the conversion efficiency of conventional boost converters, a high step-up converter is introduced as a front-end stage to stabilize the output dc voltage of each DER modules for use with the simplified multilevel inverter. Finally a Twenty Seven-level output is observed by giving three supply voltages to the multilevel inverter. To develop the model of hybrid multilevel inverter, a simulation is done based on MATLAB/SIMULINK (version 7.12) is used.

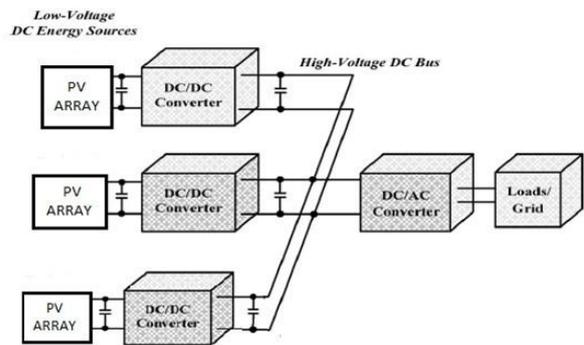


Figure 5: Configuration of DER Based Single-Phase Asymmetrical 27 Level Inverter Topology

VII. SIMULATIONS RESULTS

To verify the feasibility of the DER based single-phase asymmetrical 27-level inverter, a widely used software program MATLAB-Simulink is applied to simulate the circuit according to the previously mentioned operation principle. Input sources, the output voltages of each pv array are 30v, 90v, 270v respectively are connected as shown in fig 6 to the inverter followed a linear resistive load through the high step-up dc/dc converters. High step-up converter topology is used to boost and stabilize the

output dc voltage of DER such as various PV arrays for employment of the proposed simplified multilevel inverter. The three input voltage sources feeding from the high step-up converter is controlled at $V_{dc1} = 100v$, $V_{dc2} = 3 V_{dc1}$, $V_{dc3} = 9 V_{dc1}$. The proposed prototype of 27 level inverter is shown in fig 7 and corresponding lower inverter generates a fundamental output voltage of 1280V using three individual DC sources. The harmonic spectrum is as shown fig 9. Based on the simulation results, the proposed multilevel inverter was tested by a prototype. Components comparison between conventional inverters and the proposed approach appearing 27 output levels is given in Table 2. It is compared with the conventional multilevel inverters, i.e., diode-clamped, flying capacitor, cascaded H-bridge, and cascaded transformer based multilevel inverter. In the case of diode-clamped, a large number of clamping diodes are a severe drawback. And a lot of balancing capacitors is a disadvantage of the flying capacitor method. Among them, the isolated CML looks very effective to synthesize output voltage levels. It only needs a single dc input source. However, it shows low. Efficiency because of adopting a cascaded transformer. And it will be suffered from large size and heavy weight. Moreover, this method is not desirable for the motor drives employing VF (variable frequency) control scheme because of the saturation of transformer.

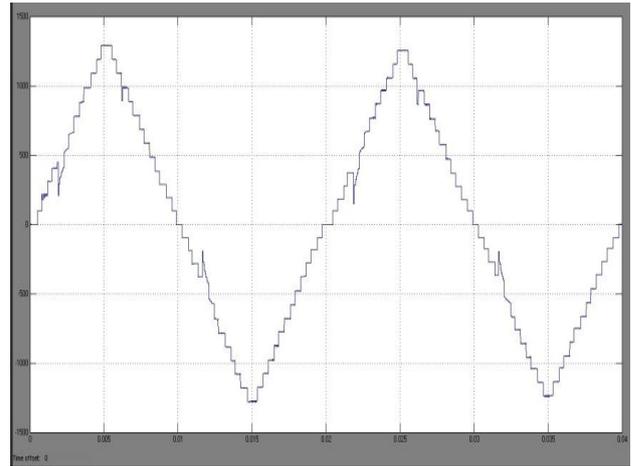


Figure 8: Twenty Seven level output

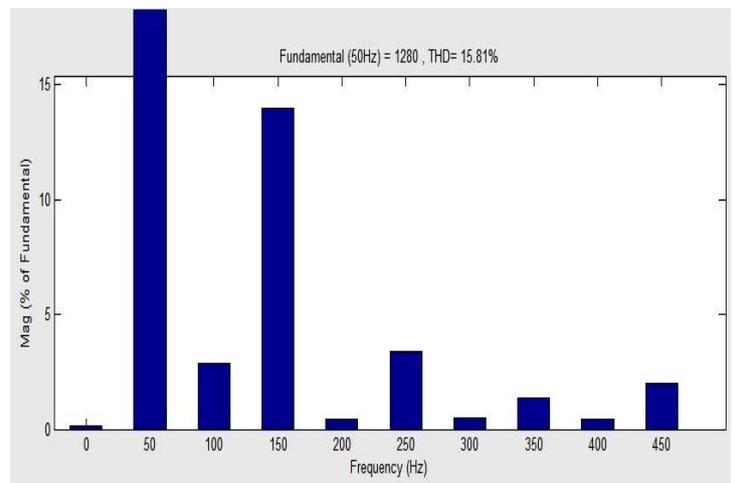


Figure 9: FFT Analysis of Twenty Seven level output

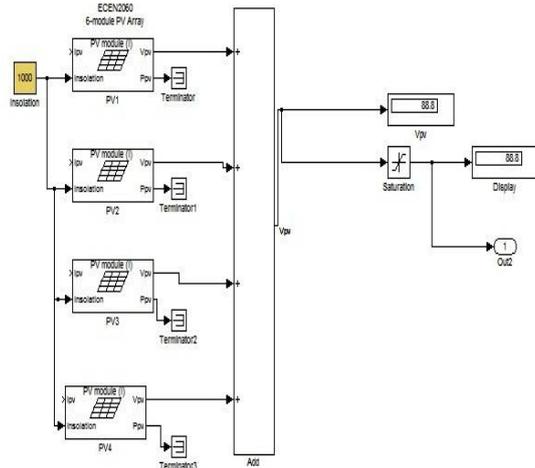


Figure 6: Simulation of Pv Array

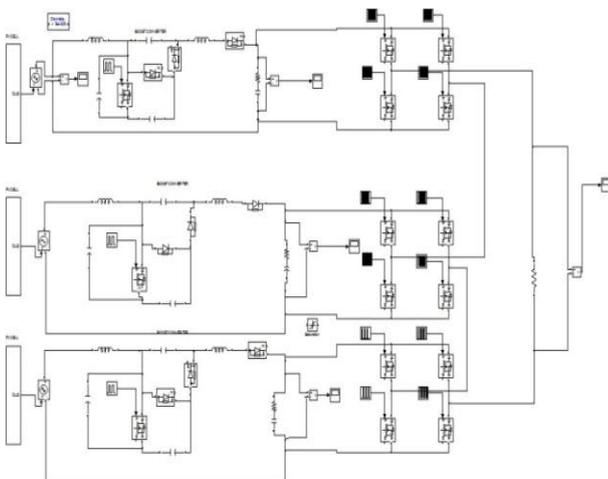


Figure 7: Simulation of Main Circuit

	Cascade	Hybrid	Proposed Prototype
No of level $S = 3$	$2S+1, 7$ level	$2^{S+1} - 1, 15$ level	$3^S, 27$ level
Input DC Voltage	$V_{dc}, 1 V_{dc}$	$2^{S-1} V_{dc}, 4 V_{dc}$	$3^{S-1} V_{dc}, 9 V_{dc}$

Table.2 Comparison of different Topologies

From the comparison, it is clear that the most outstanding advantage of the proposed multilevel inverter scheme is the elimination of transformer in the main power stage. However, each cell of the proposed multilevel inverter requires its own isolated power supply. The provision of these isolated supplies is the main limitation in the power electronic circuit design. So the proposed multilevel inverter is suitable for photovoltaic power generating systems equipped with distributed power sources.

VIII. CONCLUSION

In this Paper much effort has been focused on the development of environmentally friendly distributed energy resources (DERs) along with cascaded H-bridge multilevel inverter employing trinary dc sources to obtain a large number of output voltage levels with minimum devices.

The proposed inverter can synthesize high quality output voltage near to sinusoidal waves. The circuit configuration is simple and easy to control. The proposed prototype consists of three dc sources with the use of 12 switches. Valuable and presentable merits of the proposed approach are summarized as

- (1) Economical circuit configuration to produce multilevel outputs by using trinary input sources,
- (2) Easy to increase of the output voltage levels and output power owing to modularity characteristic,
- (3) Little transition loss of switches due to low switching frequency and reduced EMI; it is suitable for high voltage applications.

REFERENCES

- [1] Yi-Hung Liao, and Ching-Ming Lai, "Newly-Constructed Simplified Single-Phase Multistring Multilevel Inverter Topology for Distributed Energy Resources" *IEEE Transactions on Power Electronics*, Vol. 26, No. 9, September 2011
- [2] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," *IEEE Power Energy Mag.*, vol. 5, no. 4, pp. 78–94, Jul./Aug. 2007.
- [3] F. Katiraei, R. Iravani, N. Hatziargyriou, and A. Dimeas, "Microgrids management," *IEEE Power Energy Mag.*, vol. 6, no. 3, pp. 54–65, May/June 2008.
- [4] C. L. Chen, Y. Wang, J. S. Lai, Y. S. Lee, and D. Martin, "Design of parallel inverters for smooth mode transfer microgrid applications," *IEEE Trans. Power Electron.*, vol. 25, no. 1, pp. 6–15, Jan. 2010.
- [5] L. G. Franquelo, J. Rodriguez, S. Kouro, R. Portillo, and M. A. M. Prats, "The age of multilevel converter arrives," *IEEE Ind. Electron Magazine*, pp. 28–39, 2008.
- [6] J. Rodriguez, J. S. Lai, and F. Z. Peng, "Multilevel Inverters: A survey of topologies, controls, and applications," *IEEE Trans. Ind. Electron.*, vol. 49, no. 4, pp. 724–738, Aug. 2002.
- [7] J. S. Lai, and F. Z. Peng, "Multilevel Converters—A New Breed of Power Converters," *IEEE Trans. Ind. Appl.*, vol. 32, no. 3, pp. 509–517, May/June, 1996.
- [8] M. H. Rashid, *Power Electronics Handbook*, Academic Press, 2001, pp. 539–562.
- [9] L. M. Tolbert, F. Z. Peng, and T. G. Habetler, "Multilevel converters for large electric drives," *IEEE Trans. Ind. Electron.*, vol. 35, pp. 36–44, 1999.
- [10] F. S. Kang, S. J. Park, M. H. Lee, C. U. Kim, "An efficient multilevel synthesis approach and its application to a 27-level inverter," *IEEE Trans. Ind. Electron.*, vol. 52, no. 6, pp. 1600–1606, 2005.
- [11] J. P. Benner and L. Kazmerski, "Photovoltaics gaining greater visibility," *IEEE Spectr.*, vol. 29, no. 9, pp. 34–42, Sep. 1999.
- [12] E. Bezzel, H. Lauritzen, and S. Wedel. (2004) The photo electro chemical solar cell. PEC Solar Cell Project, Danish Technological Institute. [Online].
- [13] H. Wilk, D. Ruoss, and P. Toggweiler. (2002) Innovative electrical concepts. International Energy Agency Photovoltaic Power Systems, IEA PVPS 7-07:2002. [Online]. Available: www.iea-pvps.org
- [14] M. Wuest, P. Toggweiler, and J. Riatsch, "Single cell converter system (SCCS)," in *Proc. 1st IEEE WCPEC*, vol. 1, 1994, pp. 813–815.
- [15] J. Riatsch, H. Stemmler, and R. Schmidt, "Single cell module integrated converter system for photovoltaic energy generation," in *Proc. EPE'97*, vol. 1, Trondheim, Norway, 1997, pp. 71–77.
- [16] S. B. Kjaer, "Design and control of an inverter for photovoltaic applications," Ph.D. dissertation, Inst. Energy Technol., Aalborg University, Aalborg East, Denmark, 2004/2005.
- [17] W. Yu, C. Hutchens, J. S. Lai, J. Zhang, G. Lisi, A. Djabbari, G. Smith, and T. Hegarty, "High efficiency converter with charge pump and coupled inductor for wide input photovoltaic AC module applications," in *Proc. IEEE Energy Convers. Congr. Expo.*, 2009, pp. 3895–3900.
- [18] A. Mansoor, W. M. Grady, P. T. Staats, R. S. Thallam, M. T. Doyle, and M. J. Samotyj, "Predicting the net harmonic currents produced by large numbers of distributed single-phase computer loads," *Inst. Elect. Eng. Trans. Power Dist.*, vol. 10, no. 4, pp. 2001–2006, 1995.
- [19] E. F. El-Saadany and M. M. A. Salama, "Reduction of the net harmonic current by single-phase nonlinear loads due to attenuation and diversity effects," *Int. J. Elec. Power Energy Syst.*, vol. 20, no. 4, pp. 259–268, 1998.
- [20] D. G. Infield, "Combined switching harmonics from multiple grid-connected single-phase inverters," *Proc. Inst. Elect. Eng., Gen. Transm. Dist.*, vol. 148, no. 5, pp. 427–430, 2001.
- [21] S. Vazquez, J. I. Leon, J. M. Carrasco, L. G. Franquelo, E. Galvan, M. Reyes, J. A. Sanchez, and E. Dominguez, "Analysis of the power balance in the cells of a multilevel cascaded H-bridge converter," *IEEE Trans. Ind. Electron.*, vol. 57, no. 7, pp. 2287–2296, Jul. 2010.

Series of Information Divergence Measures Using New F-Divergences, Convex Properties and Inequalities

K. C. Jain^{*}, Ram Naresh Saraswat

Department of Mathematics, Malaviya National Institute of Technology, Jaipur (Rajasthan) -302017, INDIA

ABSTRACT: There are several types of information divergence measures studied in literature of information theory which compare two probability distributions and have applications in information theory, statistics and engineering. In this paper, we derive some families of divergence measures using properties of convex functions and new f-divergence measure. Bounds of new divergences in terms of well-known divergence measure are also considered.

Additional Key words and Phrases: Triangular discrimination, symmetric divergence measure, Csiszar, s f-divergence etc.

AMS Classification 62B-10, 94A-17, 26D15

I. INTRODUCTION

Let

$$\Gamma_n = \left\{ P = (p_1, p_2, \dots, p_n) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\}, n \geq 2 \tag{1.1}$$

be the set of all complete finite discrete probability distributions. There are many information and divergence measures exists in the literature of information theory and statistics. Csiszar [1] & [2] introduced a generalized measure of information using f-divergence measure is given by

$$I_f(P, Q) = \sum_{i=1}^n q_i f\left(\frac{p_i}{q_i}\right) \tag{1.2}$$

where $f: \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is a convex function and $P, Q \in \Gamma_n$.

The Csiszar's f-divergence is a general class of divergence measures that includes several divergences used in measuring the distance or affinity between two probability distributions. This class is introduced by using a convex function f , defined on $(0, \infty)$. An important property of this divergence is that many known divergences can be obtained from this measure by appropriately defining the convex function f . These measures have been applied in a variety of fields such as anthropology, genetics, finance, economics, analysis of contingency tables, approximations of probability distributions, signal processing & pattern recognition.

II. NEW F-DIVERGENCE MEASURE OF INFORMATION

Now, we shall consider some properties of a new f-divergence measure and its particular cases which are may be interesting in areas of information theory. Jain and Saraswat, [5, 6] introduced a new f-divergence measure is given by

$$S_f(P, Q) = \sum_{i=1}^n q_i f\left(\frac{p_i + q_i}{2q_i}\right) \tag{2.1}$$

Where $f: \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is a convex function and $P, Q \in \Gamma_n$.

It is shown that using new f-divergence measure we derive some well known divergence measures such as Chi-square divergence, Triangular discrimination and variational distance in this section. Other well-known divergence measures are also derived in [6] using new f-divergence measure. We derived some new series of divergence measures using properties of convex functions in section 3. Using additional properties of convex functions we derived some divergence measure of information in section 4. Some bounds of other new divergence measure in terms of various well-known divergence measures like as Triangular discrimination, Hellinger discrimination and variational distance, in section 5 are also considered.

The following results are also presented by Jain and Saraswat [6].

Proposition 2.1 Let $f: [0, \infty) \rightarrow \mathbf{R}$ be convex and $P, Q \in \Gamma_n$ then we have the following inequality

$$S_f(P, Q) \geq f(1) \tag{2.1}$$

Equality holds in (2.1) iff

$$p_i = q_i \quad \forall i = 1, 2, \dots, n \tag{2.2}$$

Corollary 2.1.1 (Non-negativity of new f-divergence measure) Let $f : [0, \infty) \rightarrow \mathbf{R}$ be convex and normalized, i.e.

$$f(1) = 0 \tag{2.3}$$

Then for any $P, Q \in \Gamma_n$ from (2.1) of proposition 2.1 and (2.3), we have the inequality

$$S_f(P, Q) \geq 0 \tag{2.4}$$

If f is strictly convex, equality holds in (2.5) iff

$$p_i = q_i \quad \forall i \in [1, 2, \dots, n] \tag{2.5}$$

and

$$S_f(P, Q) \geq 0 \quad \text{and} \quad S_f(P, Q) = 0 \quad \text{iff} \quad P = Q \tag{2.6}$$

Proposition 2.2 Let f_1 & f_2 are two convex functions and $g = a f_1 + b f_2$ then $S_g(P, Q) = a S_{f_1}(P, Q) + b S_{f_2}(P, Q)$,

where a & b are constants and $P, Q \in \Gamma_n$

We now give some examples of well known information divergence measures which are obtained from new f-divergence measure.

• **Chi-square divergence measure** [9]: - If $f(t) = 4(t-1)^2$ then Chi-square divergence measure is given by

$$S_f(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^2}{q_i} = \left[\sum_{i=1}^n \frac{p_i^2}{q_i} - 1 \right] = \chi^2(P, Q) \tag{2.7}$$

• **Triangular discrimination** [4]: - If $f(t) = \frac{2(t-1)^2}{t}$, $\forall t > 0$ then Triangular discrimination is given by

$$S_f(P, Q) = 2[1 - W(P, Q)] = \sum_{i=1}^n \frac{(p_i - q_i)^2}{(p_i + q_i)} = \Delta(P, Q) \tag{2.8}$$

where $W(P, Q) = \sum_{i=1}^n \frac{2p_i q_i}{p_i + q_i}$ is known as harmonic mean divergence measure

• **Variational Distance** [4, 8]:- Let if $f(t) = 2|t-1|$ then Chi-square divergence measure is given by

$$S_f(P, Q) = \sum_{i=1}^n |p_i - q_i| = V(P, Q) \tag{2.9}$$

NEW INFORMATION DIVERGENCE MEASURES

In this section we shall find out the new information measures with help of the following convex function. Now we consider the function $f : (0, \infty) \rightarrow \mathbf{R}$ given by

$$f_k(t) = t \left(1 - \frac{1}{t} \right)^{2k}, \quad k = 1, 2, 3, \dots \tag{3.1}$$

Since

$$f_k'(t) = \left(1 - \frac{1}{t} \right)^{2k-1} \left[\frac{(2k-1)+t}{t} \right] \tag{3.2}$$

$$\text{and } f_k''(t) = \frac{(t-1)^{2k-2} [2k(2k-1)]}{t^{2k+1}} \tag{3.3}$$

Function $f_k(t)$ is always convex if $k = 1, 2, 3, \dots$ positive integers, $\forall t > 0$.

$$S_f(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^{2k}}{(p_i + q_i)^{2k-1}} = N_k^c(P, Q)$$

Also we get the following

$$F = \sum_{i=1}^n f_k(t) = t \sum_{i=1}^{\infty} \left(1 - \frac{1}{t}\right)^{2k} = t \left(1 - \frac{1}{t}\right)^2 \left[1 + \left(1 - \frac{1}{t}\right)^2 + \dots + \infty\right] = \frac{(t-1)^2}{t} \left[1 + \left(1 - \frac{1}{t}\right)^2 + \dots + \infty\right]$$

Sum always convex for infinity if $\forall t > \frac{1}{2}$

Now we have the following series of convex functions if $k=1, 2, 3, 4, \dots$ of (3.1)

$$\frac{(t-1)^2}{t}, \frac{(t-1)^4}{t^3}, \frac{(t-1)^6}{t^5}, \frac{(t-1)^8}{t^7}, \dots$$

Further we know that if $f_1(t), f_2(t), f_3(t), f_4(t), \dots$ are convex functions then the function

$c_1 f_1(t) + c_2 f_2(t) + c_3 f_3(t) + c_4 f_4(t) + \dots$ is also convex functions. Where $c_1, c_2, c_3, c_4, \dots$ are positive constants such that at least one c_i is not equal to zero. Now taking

$$c_1 = 1, c_2 = 1, c_3 = \frac{1}{2!}, c_4 = \frac{1}{3!}, \dots$$

$$\text{and } f_1(t) = \frac{(t-1)^2}{t}, f_2(t) = \frac{(t-1)^4}{t^3}, f_3(t) = \frac{(t-1)^6}{t^5}, f_4(t) = \frac{(t-1)^8}{t^7}, \dots$$

We have the following series of convex functions

$$\frac{(t-1)^2}{t} + \frac{(t-1)^4}{t^3} + \frac{(t-1)^6}{2!t^5} + \frac{(t-1)^8}{3!t^7} + \dots$$

$$= \frac{(t-1)^2}{t} \left[1 + \frac{(t-1)^2}{t^2} + \frac{(t-1)^4}{2!t^4} + \frac{(t-1)^6}{3!t^6} + \dots\right]$$

$$g(t) = \frac{(t-1)^2}{t} \exp\left\{\frac{(t-1)^2}{t^2}\right\} \tag{3.4}$$

where $\exp\{\cdot\}$ denotes the exponential function and applying in following function new f-divergence property of (2.2)

$$S_f(P, Q) = g(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + q_i} \exp\left\{\frac{(p_i - q_i)^2}{(p_i + q_i)^2}\right\} \tag{3.5}$$

$$g(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + q_i} \exp\left\{\frac{(p_i - q_i)^2}{2(p_i + q_i)(p_i + q_i)}\right\} \tag{3.6}$$

Divergence measure $g(P, Q)$ is the combination of triangular and arithmetic divergence measures, similarly, we get

$$c_1 = 1, c_2 = 1, c_3 = \frac{1}{2!}, c_4 = \frac{1}{3!}, \dots$$

$$\text{and } f_1(t) = \frac{(t-1)^4}{t^3}, f_2(t) = \frac{(t-1)^6}{t^5}, f_3(t) = \frac{(t-1)^8}{t^7}, f_4(t) = \frac{(t-1)^{10}}{t^9}, \dots$$

then we obtain the following divergence measure of Csiszar's f-divergence class

$$= \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^4}{(p_i + q_i)^3} \exp\left\{\frac{(p_i - q_i)^2}{(p_i + q_i)^2}\right\} \tag{3.7}$$

Similarly an appropriate selection of constants and convex functions will result in following series of convex functions

$$f_k^*(t) = \frac{(t-1)^{2k}}{t^{2k-1}} \exp\left\{\frac{(t-1)^2}{t^2}\right\}, \quad k = 1, 2, 3, 4, \dots \tag{3.8}$$

and the following series of divergence measures of Csiszar's f-divergence class

$$N_k^*(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^{2k}}{(p_i + q_i)^{2k-1}} \exp\left\{\frac{(p_i - q_i)^2}{(p_i + q_i)^2}\right\} \tag{3.9}$$

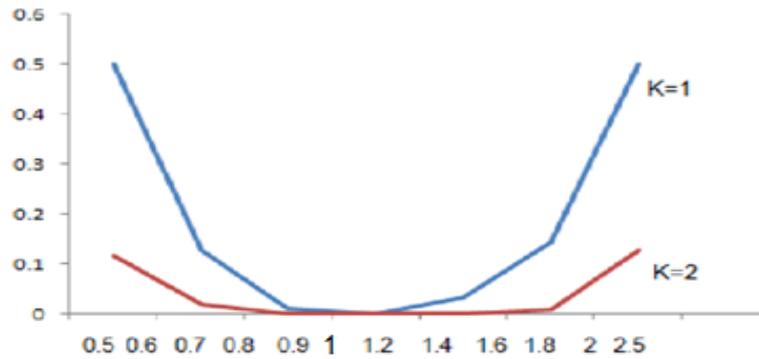


Fig.3.1 graph of the convex function $f_k(t)$

It is clear that from the above graph that the convex functions $f_k^*(t)$ gives a steeper slope with increase in value of k. Further $f_k(1) = 0$, so that $N_k^*(P, P) = 0$ and the convexity of the function $f_k(t)$ ensure that the measure (3.3) is non-negative. Thus we can say that the measure (3.3) non-negative and convex in the pair of probability distributions $(P, Q) \in \Gamma_n$. Since

$$N_1^*(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2}{(p_i + q_i)} \exp \left\{ \frac{(p_i - q_i)^2}{(p_i + q_i)^2} \right\} \quad (3.10)$$

$$N_2^*(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^4}{(p_i + q_i)^3} \exp \left\{ \frac{(p_i - q_i)^2}{(p_i + q_i)^2} \right\} \quad (3.11)$$

$$N_3^*(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^6}{(p_i + q_i)^5} \exp \left\{ \frac{(p_i - q_i)^2}{(p_i + q_i)^2} \right\} \quad (3.12)$$

....and so on.

Therefore we can say that the measure $N_k^*(P, Q)$ is made up generalized series of combinations of triangular and arithmetic divergence measure.

IV. SOME NEW OTHERS INFORMATION DIVERGENCE MEASURES

In this section, we will derive some new information divergence measures using the convexity property of divergence measures. We proceed as follows.

Since the sum of two convex functions is again a convex function, therefore we have the following convex functions.

$$\frac{(t-1)^2}{t} + \frac{(t-1)^4}{t^3} = \frac{(t-1)^2(2t^2 - 2t + 1)}{t^3} \quad (4.1)$$

$$\frac{(t-1)^4}{t^3} + \frac{(t-1)^6}{t^5} = \frac{(t-1)^4(2t^2 - 2t + 1)}{t^5} \quad (4.2)$$

$$\frac{(t-1)^6}{t^5} + \frac{(t-1)^8}{t^7} = \frac{(t-1)^6(2t^2 - 2t + 1)}{t^7} \quad (4.3)$$

..and so on.

The following divergence measures are obtained using new f-divergence measure is given by

$$N_1(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^2 (p_i^2 + q_i^2)}{(p_i + q_i)^3} \quad (4.4)$$

$$N_2(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^4 (p_i^2 + q_i^2)}{(p_i + q_i)^5} \quad (4.5)$$

$$N_3(P, Q) = \frac{1}{2} \sum_{i=1}^n \frac{(p_i - q_i)^6 (p_i^2 + q_i^2)}{(p_i + q_i)^7} \quad (4.6)$$

....and so on.

Similarly we can generate various other series of divergence measures using the properties of convex functions. Further results about these divergence measures will be discussed elsewhere.

V. RELATIONS AMONG NEW INFORMATION DIVERGENCE AND OTHER WELL KNOWN DIVERGENCE MEASURES

In this section we will drive inequalities relating $N_k^*(P, Q)$ (for the case $k=1, k=2, \dots$) with the above divergence measures.

To start with, we will derive inequalities for the divergence measures given by $\Delta_1(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^2}{(p_i + q_i)}$,

$$\Delta_2(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^4}{(p_i + q_i)^3}, \dots, \Delta_k(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^{k+1}}{(p_i + q_i)^k} \quad (5.1)$$

The above divergence measures (5.1) may be say the generalized triangular divergence measure of type k. Then we will use these inequalities for relating $N_k^*(P, Q)$ with other divergence measure.

Now again consider the inequality

$$x + 1 \leq \exp\{x\}$$

Replacing x by $\frac{(t-1)^2}{t^2}$

$$\frac{(t-1)^2}{t^2} + 1 \leq \exp\left\{\frac{(t-1)^2}{t^2}\right\}$$

$$\frac{2t^2 - 2t + 1}{t^2} \leq \exp\left\{\frac{(t-1)^2}{t^2}\right\}$$

$$\frac{(t-1)^2(2t^2 - 2t + 1)}{t^3} = \frac{(t-1)^2}{t} + \frac{(t-1)^4}{t^3} \leq \frac{(t-1)^2}{t} \exp\left\{\frac{(t-1)^2}{t^2}\right\}$$

$$\frac{(t-1)^2}{t} + \frac{(t-1)^4}{t^3} \leq \frac{(t-1)^2}{t} \exp\left\{\frac{(t-1)^2}{t^2}\right\}$$

$$N_1(P, Q) = \sum_{i=1}^n \frac{(p_i - q_i)^2(p_i^2 + q_i^2)}{(p_i + q_i)^3} \leq N_1^*(P, Q) \quad (5.2)$$

Now replacing t by $\frac{p+q}{2q}$, multiplying q and finally summing over all t in the above inequality, we obtain

$$N_1(P, Q) \leq N_1^*(P, Q) \quad (5.3)$$

$$\Delta(P, Q) \leq N_1(P, Q) \leq N_1^*(P, Q) \quad (5.4)$$

$$\therefore \Delta(P, Q) \leq \chi_1^2(Q, P) \quad [\text{Jain \& Saraswat, 5}]$$

$$1/2V^2(P, Q) \leq \Delta(P, Q) \leq N_1(P, Q) \leq N_1^*(P, Q) \quad (5.5)$$

$$\therefore \frac{1}{2}V^2(P, Q) \leq \Delta(P, Q) \quad [\text{Topse, 9}]$$

$$2h(P, Q) \leq \Delta(P, Q) \leq N_1(P, Q) \leq N_1^*(P, Q) \quad (5.6)$$

$$\therefore 2h(P, Q) \leq \Delta(P, Q) \quad [\text{Dacunha-Castelle, 3}] \text{ where } h(P, Q) \text{ is Hellinger discrimination [4]}$$

REFERENCES

- [1] **Csiszar I. Information measure**, A critical survey. Trans. 7th Prague conf. on info. Th. Statist. Decius. Funct, Random Processes and 8th European meeting of statist Volume B. Acadmia Prague, 1978, PP-73-86
- [2] **Csiszar I.**, Information-type measures of difference of probability functions and indirect observations. *studia Sci.Math.hunger.*2(1961).299-318
- [3] **Dacunha D.-Castelle**, *Ecole de Probabilites de Saint-Flour VII-1977*, Berlin, Heidelberg, New York: Springer, 1978
- [4] **Hellinger E.**, Neue Begrundung der Theorie der quadratischen Formen Von unendlichen vielen veranderlichen, *J.Rein.Aug.Math*,136(1909),210-271.
- [5] **Jain K. C. and R. N. Saraswat** "A New Information Inequality and its Application in Establishing Relation among various f-Divergence Measures", *Journal of Applied Mathematics, Statistics and Informatics (JAMSI)*, 8(1) (2012), 17-32.
- [6] **Jain K. C. and R. N. Saraswat** "Some bounds of information divergence measure in term of Relative arithmetic-geometric divergence measure" *International Journal of Applied Mathematics and Statistics*, Volume 32, Number 2 processing (2013), pp. 48-58.
- [7] **Kolmogorov, A. N.**, A new invariant for transitive dynamical system. *Dokl. Akad. Nauk USSR.* (1957) 119, 861-869
- [8] **Pearson K.**, On the criterion that a give system of deviations from the probable in the case of correlated system of variables in such that it can be reasonable supposed to have arisen from random sampling, *Phil. Mag.*, 50(1900),157-172.
- [9] **Topse F.** Some inequalities for information divergence and related measure of discrimination. *Res. Rep. Collection, RGMIA*,2,1,85-98

Wear behavior of GI250 grade of grey cast iron coated by WC-12CO and Stellite-6 coatings

Baljit Singh^{1*}, Neel Kanth Grover², Jwala Parshad Singla²

¹Department of mechanical engg DBFGOI, Moga.

²Department of mechanical engg SBSCET Ferozepur.

²Department of mechanical engg SBSCET Ferozepur.

ABSTRACT: The wear is very severe problem in industry. It leads economic loss to society. In the present research study efforts will be made to reduce the wear rate of Braking Disc Rotor with surface modifications techniques and comparison of wear properties grade of cast iron GI250 coated by WC-12CO and Stellite-6 deposited by Detonation spray processes is presented. The results show that the WC-12CO on GI250 grey cast iron performs slightly better than the Stellite-6 coating. The WC-12CO-GI250 coating substrate combination has shown minimum Cumulative Volume loss among all the two combinations. The wear resistance for coating-substrate combinations in their decreasing order is WC-12CO-GI250 > Stellite-6-GI250.

Keywords: Wear, Detonation spray, Brake disc wear resistance.

I. INTRODUCTION

The optimization of automotive vehicles braking systems, subjected to mechanical and thermal stresses, depends on a

Table 2.1 Chemical composition (Wt %) of the GI250 grey iron

	C	Si	Mn	P	S	Cr	Mo	Ni	Cu
GI250	3.35	2.16	0.53	0.118	0.127	0.658	0.025	0.293	0.250

Table 2.2 Hardness values of GI250 grey iron

Material	GI250
Hardness (HB)	285

2.3. DEPOSITION OF COATINGS

2.3.1 Detonation Spray Coating Powders

Two types of coating powders namely (1) Stellite-6 (2) WC-12CO are selected for Detonation Spray Coating Process after the literature survey. Research papers shows that the above coatings have excellent wear resistance.

(1) Stellite-6: Stellite powder when sprayed using the Awaaz Detonation spray coating process, it produces coatings which are very dense and homogeneous. The various constituents of the powder are Carbon (1%), Tungsten (6 %), Fe (5%), Chromium (30%), and Cobalt (58%). The powder particle size is 10-45 microns.

(2) Tungsten carbide (WC-12CO): WC-12CO when sprayed using the Awaaz Detonation spray coating process, it produces coatings, which are very hard, dense and

combination of properties. In general, a complex state of stress is found and it is practically impossible to select a material and design a component based only on one of these properties. The material used in brake rotors should be able to bear thermal fatigue and should absorb and dissipate, as soon as possible, the heat generated during braking [1].

2.2. PREPARATION OF SAMPLES

Small cylindrical pins having diameter of circular cross-section equal to 8mm and length equal to 30 mm were prepared from GI250 grey irons. A total of 9 pins of GI250 grade were prepared. Pins were given the Sample No. from 1 to 9. The grinding of end faces (to be coated) of the pins done using emery papers of five different grades 220, 400, 600, 800, 1000 in the same order . Grinding was followed by polishing with 1/0, 2/0, 3/0 and 4/0 grades polishing papers.

excellent bonded. Coatings can be built up to higher thickness. WC-12CO when sprayed using the detonation spray coating process, it produces coatings, which are very hard, dense and its bonding strength is very high. Coatings can be built up to higher thickness. These are coatings which are gives better protection against wear. The powder particle size is 15-35 microns. The various constituents of the powder are Chromium (32.72%), Cobalt (46.58%) and Tungsten (20.70%).

Table 2.3 Detonation Spray parameters for two coatings (a) Stellite-6

Gases	Oxygen	Acetylene	Nitrogen
Flow Rate	2990	2410	720

Spraying Distance 165 mm

(b) WC-12CO

Gases	Oxygen	Acetylene	Nitrogen
Flow Rate	4350	2300	960

Spraying Distance 150 mm

2.4 CHARACTERIZATION OF COATINGS

2.4.1 Specimen preparation

Two specimens having dimensions 20mm* 15mm* 5mm were cut from the substrate material GI250. The specimens were grinded using sand papers of 220, 400, 600 and 1000 grit sizes and subsequently polished on 1/0, 2/0, 3/0 grades. Samples were well polished until it shines like mirror. One sample of GI250 substrate was coated with Stellite-6 coating and other is coated with WC-12CO coating.

2.5 SLIDING WEAR STUDY USING PIN- ON -DISC CONFIGURATION

2.5.1. Experimental Set Up

Dry sliding wear tests for the uncoated and detonation spray coated cylindrical specimens were conducted using a pin-on-disc machine (Wear and Friction Monitor Tester TR-201 made by M/S DUCOM, Bangalore, INDIA) conforming to ASTM G 99 standard. The tests were conducted in air with a room temperature of 30-32°C. Wear tests were performed on the pin specimens that had flat surfaces in the contact regions and the rounded corner. The pin was held stationary against the counter face of a rotating disc made of carbon steel (EN-31) at 40 mm track diameter. EN-31 steel is a plain carbon steel; case hardened 62 to 65 HRC as provided with the pin-on-disc machine. The composition of the material of the steel disc is given in Table 2.4.

Table 2.4. Chemical composition (wt %) of the En-31 carbon steel disc

C	Si	Mn	S	P
0.42 (max)	0.05-0.35	0.40-0.70	0.05 (max)	0.05 (max)

2.5.2. Sliding Wear Studies

The pins were polished with emery paper and both disc and the pin were cleaned and dried before carrying out the test. The pin was loaded against the disc through a dead weight loading system. The wear tests for coated as well as uncoated specimens were conducted under three normal loads of 40 N, 50 N and 60 N and a fixed sliding velocity of 1 m/s. The track radii for the pins were kept at 40 mm. The speed of the rotation of the disc (477 rpm) for all the cases was so adjusted so as to keep the linear sliding velocity at a constant value of 1 m/s. A variation of ± 5 rpm was observed in the rpm of the disc. Wear tests have been carried out for a total sliding distance of 5400 m (6 cycles of 5min, 5min, 10min, 10min, 20min, 40min duration), so that only top coated surface was exposed for each detonation sprayed sample. Tangential force was monitored continuously during the wear tests. Weight losses for pins were measured after each cycle to determine the wear loss. The pin was removed from the holder after each run, cooled to room temperature, brushed lightly to remove loose wear debris, weighed and fixed again in exactly the same position in the holder so that the orientation of the sliding surface remains unchanged. The weight was measured by a micro balance to an accuracy of 0.0001 gm. The coefficient

of friction has been determined from the friction force and the normal loads in all the cases.

2.5.3. Wear rate

The wear rate data for the coated as well as uncoated specimens were plotted with respect to sliding distance to establish the wear kinetics. The specific wear rates for the coated and uncoated material were obtained by

$W = \delta w / L \rho F$ Where W denotes specific wear rates in, Bowden (B) ($1B = 10^{-6} \text{ mm}^3/\text{N}\cdot\text{m}$)

[Recommendation from IRG OECD meeting with about 30 participants to introduce a new unit for wear rate: Bowden (B) equal to $10^{-6} \text{ mm}^3/\text{N}\cdot\text{m}$]

δw is the weight loss measured in, g

L the sliding distance in, m

ρ the density of the worn material in g/mm^3 and

F the applied load in N.

2.5.4. Wear volume

The wear volume loss was also calculated from the weight loss and density of the coatings as well as substrate material for all the investigated cases. These data were reported in the form of plots showing the cumulative wear volume loss Vs sliding distance for all the cases. Bar charts were also drawn to show net

Volume = mass / density

Wear Volume Loss = $(\delta w / 9.81) / \rho$

Where δw is the weight loss in, g

And ρ is the density of material, g/mm^3

2.5.5. Coefficient of friction

The coefficient of friction (μ) determined from the frictional force and the normal load has been plotted against the sliding time to give the friction behavior of the coated as well as the uncoated material. The coefficient of friction (μ) was calculated as below:

$\mu = \text{Frictional Force (N)} / \text{Applied Normal Load (N)}$

III. RESULTS AND DISCUSSION

Characterization, Wear behaviour and SEM/EDAX analysis of Detonation sprayed Stellite-6 and WC-12 CO coatings deposited on grey cast irons has been described. Wear behaviour and SEM/EDAX analysis of bare grey cast irons has also been described and compared with that of coatings.

3.1 RESULTS

3.1.1 Characterization of Coatings

3.1.1.1 SEM/EDS analysis of the D-gun as sprayed coatings

The Scanning electron microscope micrographs as well as Energy Dispersive Spectrum (EDS) with element composition for Detonation sprayed Stellite-6 and WC-12CO coatings on GI250 shown in Figure (3.1). The microstructure of these coatings is hardly bonded, homogeneous and free from surface cracks, pores and voids. The SEM/EDS analysis of the stellite-6 coating showed in Figure (3.1(a)). The elements for stellite-6 coating corresponding to (spectrum 1) for load 40N and 50N are C, O, Cr, CO Si etc. The color of the surface at this spectrum is dull grey and near this point surface is white which may be due to presence of oxygen and at (spectrum

4) the coating is more uniform. The elemental composition for WC-12CO coating corresponding to spectrum 1 & 4 for load 40N and 50N is shown in (Figure 3.1(b)). The spectrum 4 of WC-12CO coating also confirms the presence of desired coating elements Chromium (32.72%), Cobalt (46.58%) and Tungsten (20.70%). At (spectrum 4) the color is naturally white which is may be due presence of excess oxygen on the surface.

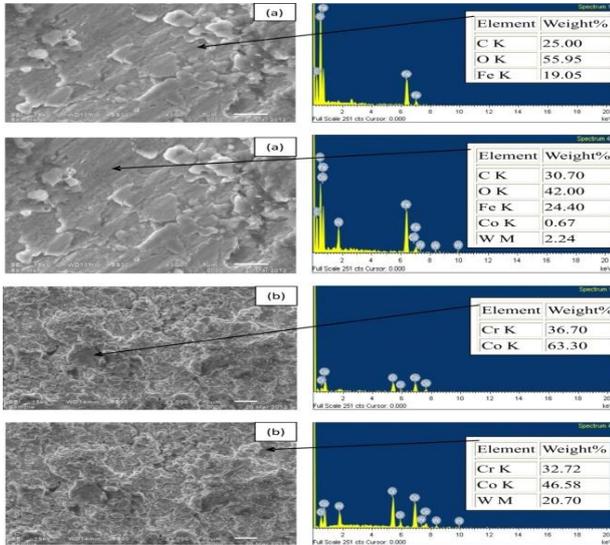


Figure 3.1 Surface Morphology and EDS patterns from different spots on as coated samples (a) Stellite-6 (b) WC-12CO.

3.1.1.2 X-ray Diffraction (XRD) analysis

The X-ray diffraction patterns for detonation sprayed stellite-6 and WC-12CO ON GI250 are shown in Figure 3.2. Figure 3.2 (a) shows the X-ray diffraction patterns for as coated samples of Stellite-6 coating on GI250 and Figure 3.2(b) shows the X-ray diffraction patterns for as coated samples of WC-12CO coating on GI250. From Figure 3.2 (a) it is identified that coating stellite-6 shows the excess of desired coating elements such as C, Cr, O, Co and small amount of Fe and Si. Similarly, from Figure 3.2(b) it is evident that coating WC-12CO shows the major phases of tungsten which is desired element of coating and minor phases of C and O which together to make Co which is also desired element of coating WC-12CO. The no. of peaks corresponding to elements of coatings can be seen from diffraction patterns of different coatings for GI250.

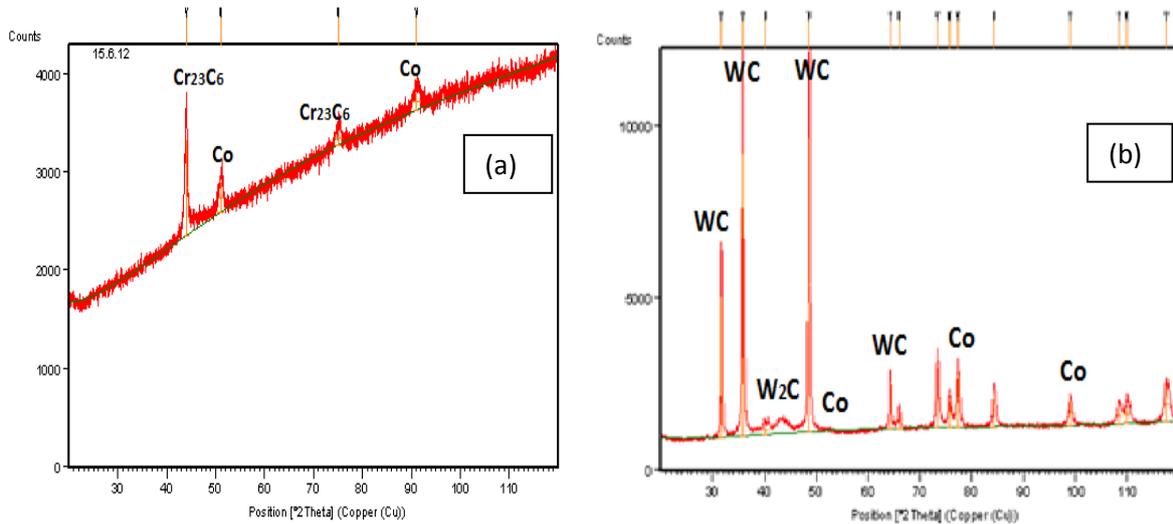


Figure 3.2 X-ray Diffraction patterns of as coated GI250 material; (a) Stellite-6 (b) WC-12CO

3.1.2 Wear Behavior

3.1.2.1. Substrate GI250

3.1.2.1.1. Wear Behavior of Two coatings vs GI250

Three samples of each coating i.e. Stellite-6 and WC-12CO on GI250 were subjected to wear on Pin-On-Disc-wear test rig at normal loads of 40N, 50N and 60N respectively. Three samples of bare GI250 substrate were also subjected to wear on Pin-On Disc-wear test rig at the same loads. The cumulative volume loss vs time for each coating is plotted as shown in Figure 3.4. From the results

of (Figure 3.4) it is investigated that cumulative volume loss for two detonation sprayed wear resistant coatings show better wear resistant in comparison to bare GI250. The bar chart (Figure 3.5) showing the Cumulative Volume Loss (CVL) in one complete cycle (90 min) is also drawn for each coating and bare GI250 substrate. From (Figure 3.4 & Figure 3.5) it is observed that with increase in load wear loss increases for the detonation sprayed coatings Stellite-6 and WC-12CO and bare GI250 the observation is same as that of the observation of (Cueva, 2003).

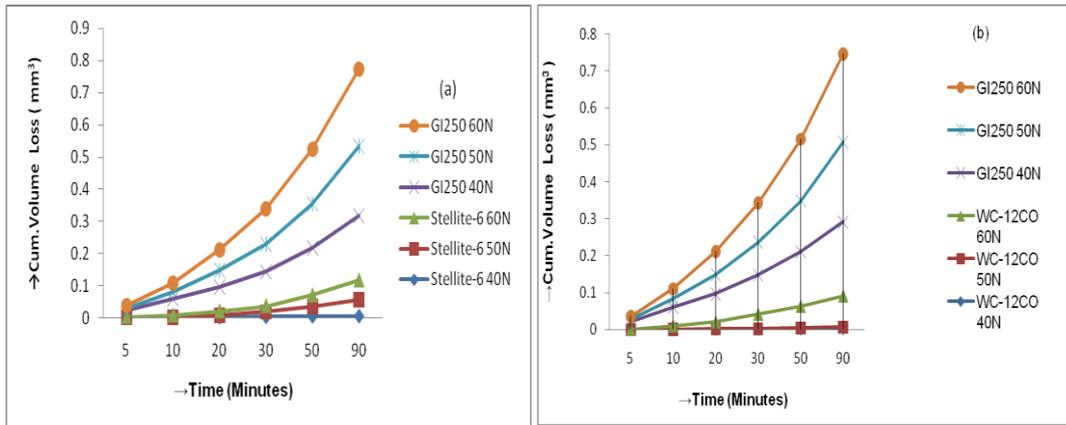


Figure 3.4 Cumulative Volume Loss (mm³) with time for (a) Stellite-6 (b) WC-12CO coatings and GI250 substrate.

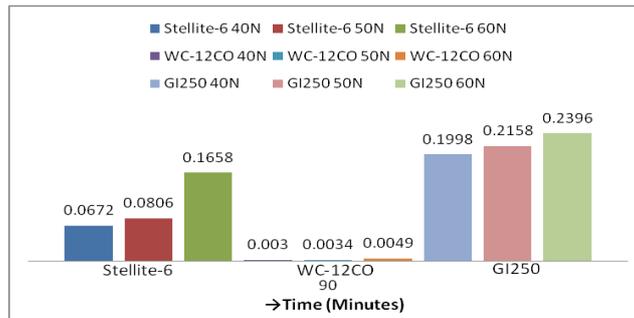


Figure 3.5 Cumulative Volume loss (mm³) in one cycle for D-gun sprayed coatings and bare GI250 at 40N, 50N and 60N.

3.1.2.1.2 Comparative Wear Behaviour for two coatings

Figure 3.6 shows Comparative Volumetric Wear Loss (mm³) for two coatings at (a) 40 N (b) 50 N and (c) 60 N. It is also observed from the results that WC-12CO is

showing the minimum cumulative volume loss as compared to other two coatings. Therefore the wear resistance of Detonation sprayed coatings on GI250 in their decreasing order can be given as WC-12CO>Stellite-6.

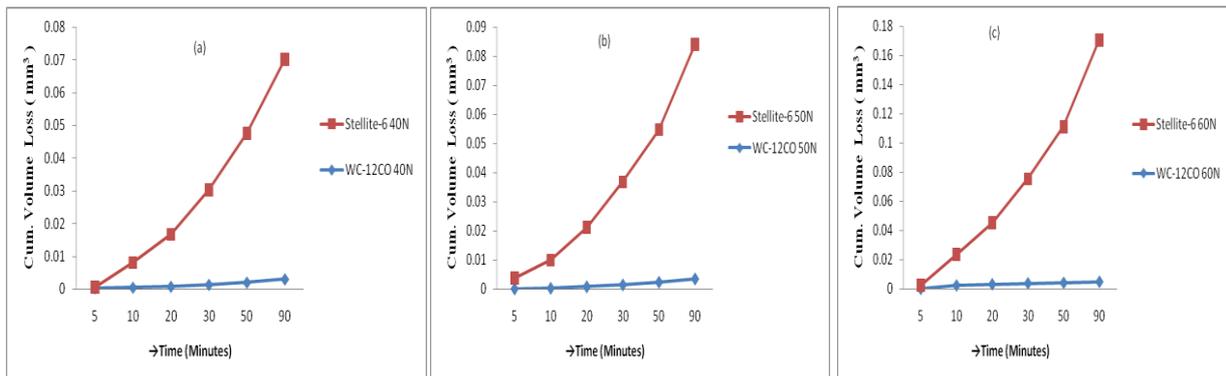


Figure 3.6 Comparative Volumetric Wear Loss (mm³) for two coatings on GI250 substrate at (a) 40 N (b) 50 N and (c) 60.

3.2 DISCUSSION

Selection of material, coating and the wear behaviour of the uncoated GI250 and detonation sprayed Stellite-6 and WC-12CO coatings have been discussed. From the study of Mohanty, 1996 it was observed that it is possible to deposit almost any material on any substrate by D-gun spray process to considerably extend the life of parts, also it is observed in the present study that the Stellite-6 and WC-12CO coatings powders have been successfully deposited on GI250 substrate by the detonation spray process. It was further confirmed by characterization of coatings using SEM and EDAX analysis of as coated specimens. SEM/EDAX results of WC-12CO and Stellite-6 (Figure 3.1 (a&b)) shows the presence of O and C on the surface which may be due to formation of carbide. Also XRD analysis of study did (Figure 3.2) which supports the results of scanning electron microscope (SEM). There is always the material loss of bare material greater than the as coated material. From (Figure 3.4) it is observed that the detonation sprayed wear resistant coatings Stellite-6, WC-12CO coated GI250 specimens showed significantly lower cumulative volume loss as compared to bare GI250 material under the normal load of 40N, 50N and 60N. It is investigated with the help of Pin-on-Disk Wear testing machine. There are many studies; Murthy & Venkataraman, 2006, Sundararajan, 2005 and Jun Wang 2000 which support the above finding that Detonation sprayed coatings increases the wear resistant and wear loss of bare material is always greater than the as coated material which is also found in present study in (Figure 3.4). Also Figure 3.6 shows comparison of two coatings in which WC-12CO has minimum wear loss as compared to Stellite-6 and therefore WC-12CO can be used for coating the grey cast iron material which is used in light truck pads. WC-12CO coating has more bonding strength than stellite-6. From (Figure 3.6) it is observed that the wear loss is increase with increase load. From this figure it is clear that the wear loss for two coatings and bare GI250 also increases with load which is same observation as Cueva (2003), The CVL for WC-12CO coating was found to be minimum in present study as shown in Figure 3.5 and Figure 3.8. it is may be due to the presence of W, CO, FE and O also tungsten and carbide increases the property of wear resistant. Identical results have been reported by Mohanty, 1996. It may be due to carbide formation due to diffusion of the Fe from the substrate. The WC-12CO-GI250 coating substrate combination has shown minimum wear loss among all four combinations as shown in Figure 3.6.

IV. CONCLUSIONS

Based upon experimental results obtained in the present study, the following conclusions have been drawn:

1. GI250 is best grade of grey cast iron.

2. Stellite-6 and WC-12CO are best coating powders to deposit on grey cast irons (GI250).
3. Stellite-6 and WC-12CO wear resistant coatings have successfully been deposited on GI250 grade of grey cast iron.
4. The Stellite-6 and WC-12CO coating based GI250 specimens showed lower cumulative volume loss as compared to bare GI250 specimens.
5. Wear loss for detonation sprayed wear resistant coatings Stellite-6 and WC-12CO coated samples and bare GI250 samples increases with increase in load.
6. The Cumulative Volume loss for WC-12CO coating was minimum in the present study. Therefore WC-12CO is best coating to deposit on GI250 grade of grey cast irons.
7. The wear resistance for coating-substrate combinations in their decreasing order is WC-12CO-GI250>Stellite-6-GI250. Therefore out of these combinations WC-12CO-GI250 coating substrate combination is the best combination.

REFERENCES

- [1] Y. Jimbo, et al., Development of high thermal conductivity cast iron for brake disk rotors, SAE Technical Paper Series, International Congress and Exposition, Detroit, MI, 1990.
- [2] H Singh, M S Grewal, H S Sekhon, and R G Rao, "Sliding wear performance of high-velocity oxy-fuel spray Al₂O₃/TiO₂ and Cr₂O₃ coatings" 29 February 2008 published Volume: 222, Issue: 4, Pages: 601-610.
- [3] Manoj Kumar Singla, Harpreet Singh, Vikas Chawla, "Thermal Sprayed CNT Reinforced Nanocomposite Coatings" Vol. 10, No.8, pp.717-726, 2011.
- [4] R. Heath, P. Heimgartner, G. Irons, R. Miller, and S. Gustafsson, Material Science Forum, Vol. [251]-54, 1997, pp 809-816.
- [5] G. Cueva, A. Sinatora, W.L. Guesser, A.P. Tschiptschin, "Wear resistance of cast irons used in brake disc rotors" Wear 255 (2003) 1256-1260.
- [6] Jun Wang, Baode Sun, Qixin Guo, Mitsuhiro Nishio and Hiroshi Ogawa, "Wear resistance of a Cr₃C₂-NiCr detonation spray coating" Volume 11, 2000, Number 2, 261-265.
- [7] J.K.N Murthy, D.S Rao, B Venkataraman, "Effect of grinding on the erosion behaviour of a WC-Co-Cr coating deposited by HVOF and detonation gun spray processes" Volume 249, Issue 7, July 2001, Pages 592-600.
- [8] M. Mohanty, R.W. Smith, M. De Bonte, L.P. Celis, E. Lugscheider, "Sliding wear behaviors of thermally sprayed 75/25 Cr₃C₂/NiCr wear resistant coatings" Wear 198 (1996) 251-266.

Evaluation of the Aquifer Hydraulic Characteristics from Electrical Sounding Data in Imo River Basin, South Eastern Nigeria: the Case of Ogwashi- Asaba Formation.

Ikechukwu Boniface Ijeh

Department of Physics, College of Natural and Applied Sciences, Michael Okpara University of Agriculture, Umudike. PMB 7267, Umuahia, Abia State, Nigeria.

ABSTRACT: Thirty two vertical electrical sounding data sets were obtained from various parts of the study area, located within latitudes $5^{\circ} 30'N$ and $5^{\circ} 45'N$, and longitudes $7^{\circ} 00'E$ and $7^{\circ} 30'E$, in the middle Imo River Basin. The litho-stratigraphic unit of Ogwashi-Asaba Formation was investigated. The sounding data were analyzed with the RESIST software to delineate the sub-surface layering. Three soundings were made at existing boreholes for comparison. The concept of Da-Zarrouk parameters (transverse unit resistance and longitudinal conductance in porous media) was used to determine aquifer hydraulic characteristics. The following values of hydraulic conductivity and transmissivity were obtained for the formation: $K_{mean}=6.71m/day$; $T_{mean}=327.91m^2/day$; $K_{max}=19.71m/day$; $T_{max}=2906.52m^2/day$; $K_{min}=1.13m/day$; $T_{min}=8.01m^2/day$. The aquifer thicknesses range from 18m to 148m for the Formation. It was established from the study that thick productive aquifers are located in the Ogwashi-Asaba Formation.

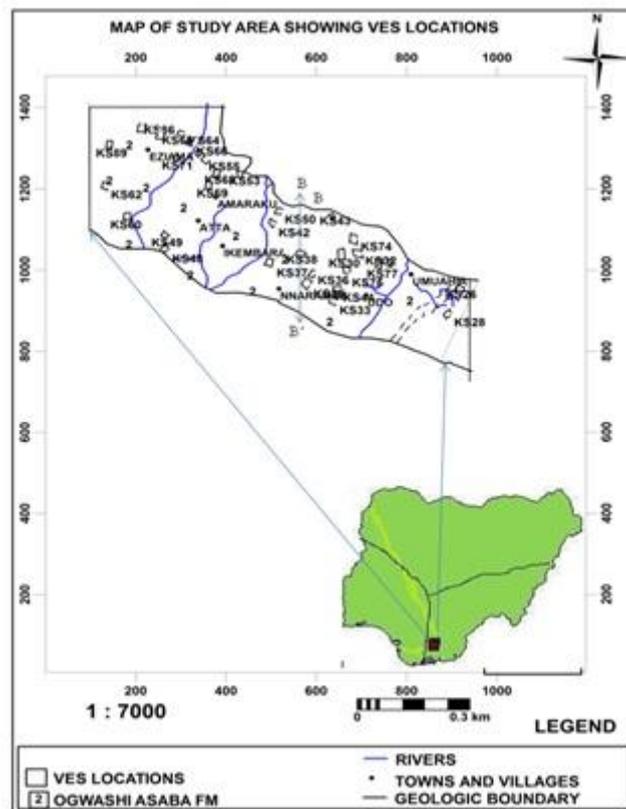
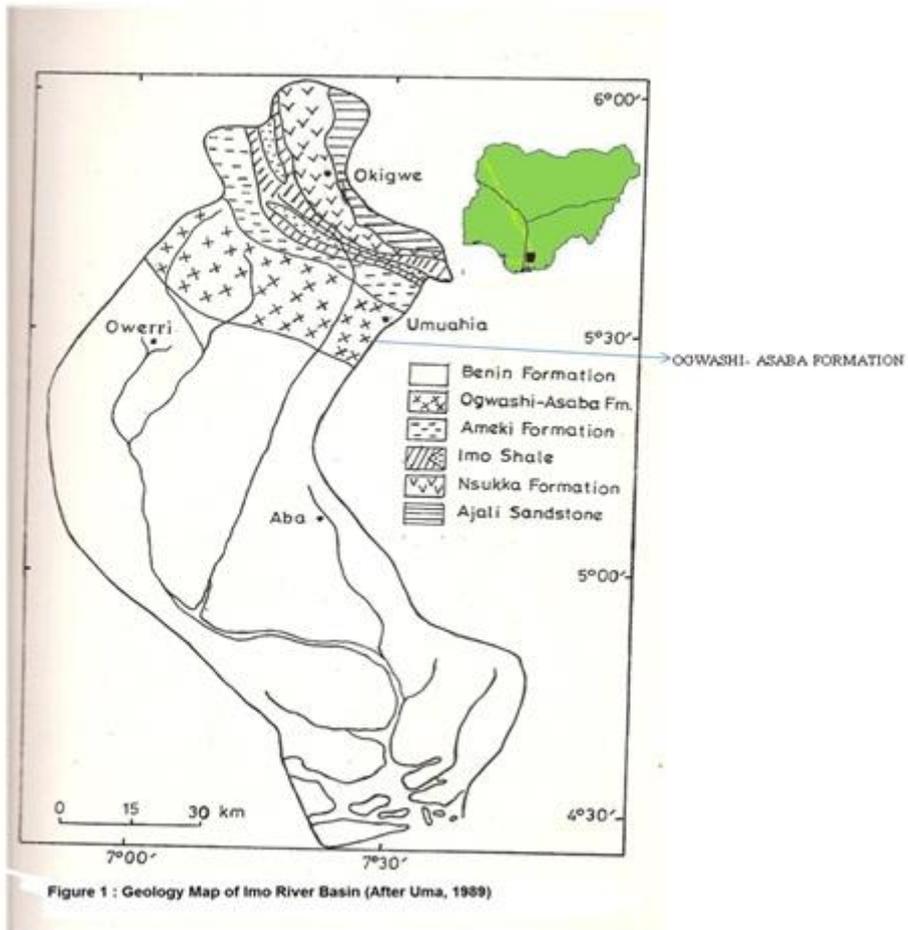
Keywords: Imo River Basin, Da-Zarrouk parameters, Aquifer hydraulic characteristics.

I. Introduction

Determination of aquifer characteristics is essential to the solution of several hydrological and hydrogeological problems. Fluid transmissivity, transverse resistance, longitudinal conductance, hydraulic conductivity and aquifer depth are very useful in describing subsurface hydrology. Much geophysical investigation of groundwater is directed towards the determination of the spatial distribution of the above mentioned aquifer hydraulic parameter (Mendoza et al., 2003). Aquifer characteristics calculated from existing boreholes are usually correlated with surface resistivity measurements based on their relationship with pore space and heterogeneity (Rubin and Hubbard, 2005; De Lima et al., 2005; Niwas et al., 2006). The Imo River Basin is based on a bedrock of a sequence of sedimentary rocks of about 5480m thick and with ages ranging from Upper Cretaceous to Recent (Uma, 1989). It is known to contain several aquiferous units. The characteristics of these aquifers such as transmissivity, hydraulic conductivity and storage potentials are not clearly understood. Since the mid 1980's, some researchers from the academia have carried out geological/geochemical investigations. Uma (1989) carried out a study on the groundwater resources of the Imo River Basin using hydro-geological data from existing boreholes. He concluded that three aquifer systems (shallow, middle and deep) exist in the area. His data were, however too sparse to make any general statement on the hydraulic characteristics of the middle Imo River Basin aquifers. Geophysical investigations on groundwater resources in the Imo River Basin were also carried out in different sections of the basin. While the contributions made by these workers are remarkable, more work still needs to be done, particularly in the area of geophysical studies, which so far have covered only a small fraction of the area of the basin. The present study is aimed at the estimation of geometry, hydraulic conductivity and transmissivity of the aquifers within the Ameki and Imo shale Formations of the Imo River Basin using the electrical resistivity method. Thirty vertical two electrical soundings (VES) were obtained at various locations within the study area.

1.1 The Study Area

Figure 1 shows the location map of the Imo River Basin where the study area is situated. The study area (Figure 2) lies between latitudes $5^{\circ}30'N$ and $5^{\circ}45'$ and longitudes $7^{\circ}00'$ and $7^{\circ}30'$. Some major communities within the study area include: Amarauku, Atta, Ikeduru, Nnarambia, Umuahia, Ogbe, Obowo and Nkwerre. The Ogwashi-Asaba formation (Oligocene to Miocene) consists in the Coastal Plain Sand which is composed of non-indurated sediments. It is generally made up of clays, sands, grits and seams of lignite alternating with gritty clay (Dessauvague, 1972). This formation is characterized by its up dip and down dip pinch outs within the Imo River Basin. The Imo River Basin has a large amount of recharge; estimated at 2.5 billion m^3 per annum, coming mainly from direct infiltration of precipitation. Average annual rainfall is about 2000mm (Onwuegbuche, 1993).



II. Methodology

Geoelectric resistivity method has been used extensively for structural, hydrological, geothermal investigations (Majumdar and Das, 2011). Geoelectrical investigations were employed to delineate formations, distinguish between sandy, shale, clay, and other layers and establish the depth to the water table, and determine the nature of the overburden. Schlumberger resistivity sounding method was used in this research. Thirty two vertical electrical soundings were carried out to establish the characteristics of the aquifers in the study area. Modeling of VES results was done using the RESIST software, which is an iterative inversion-modeling program. Analysis of the resulting apparent resistivity versus the half-current electrode separations yielded layered earth models composed of individual layers of specified thickness and apparent resistivity. The ABEM Terrameter SAS 4000, was used to obtain VES data from the field. The VES locations are shown in figure 2 coded KS.

2.1 Aquifer Hydraulic Characteristics from Vertical Electrical Sounding Data

The primary purpose of the resistivity method is to measure the potential differences on the surface due to the current flow within the ground. Since the mechanisms which control the fluid flow and electric current and conduction are generally governed by the same physical parameters and lithological attributes, the hydraulic and electric conductivities are dependent on each other.

To obtain a layer parameter, a unit square cross sectional area is cut out of a group of n-layers of infinite lateral extent. The total transverse resistance R is given by:

$$R = \sum_{i=1}^n h_i \rho_i \quad (1)$$

For a horizontal, homogeneous and isotropic medium

$$\rho = (R_1 - R_2)/(h_i - h_2) \quad (2)$$

where h_i and ρ_i are respectively the thickness and resistivity of the i^{th} layer in the section. The total longitudinal conductance S is:

$$S = \sum_{i=1}^n \frac{h_i}{\rho_i} \quad (3)$$

The longitudinal layer conductance S_i can also be expressed by

$$S_i = \sigma_i h_i \quad (4)$$

where σ_i is the layer conductivity. Conductivity in this case is analogous to the layer transmissivity, T, given by:

$$T = K_i h_i \quad (5)$$

K_i is the hydraulic conductivity of the i^{th} layer of thickness h_i . R and S of equations 1 and 3 are called the Dar Zarrouk parameters, which have been shown to be powerful interpretational aids in groundwater surveys (Zohdy et al, 1974)

According to the fundamental Darcy's law, the fluid discharge, Q, is given by

$$Q = KIA \quad (6)$$

Where K is the hydraulic conductivity, I is the hydraulic gradient, A is the cross-sectional area perpendicular to the direction of flow. Ohm's law gives:

$$j = \sigma E \quad (7)$$

where j is the current density; and σ is the electrical conductivity, which is the reciprocal of the resistivity, ρ . For aquifer material having unit cross-sectional and thickness h, the two fundamental laws can be combined to give, according to Niswam & Singhal (1981):

$$T = K \sigma R = KS/\sigma = Kh \quad (8)$$

Where T is the transmissivity; R is the transverse resistance of the aquifer, K is the hydraulic conductivity and S is the longitudinal conductance.

It has also been shown by Niwas and Singhal (1981) that in areas of similar geologic setting and water quality the product $K\sigma$ remains fairly constant. Thus, knowledge of K from some existing boreholes and of σ from VES sounding can be used to estimate $K\sigma$ for the same geologic zone. Hence, the aquifer hydraulic conductivity and transmissivity for the entire area can be estimated. This relationship forms the basis for the determination of aquifer hydraulic parameters used in this study.

III. Results and Discussion

3.1 Aquifer Characterization of the Study Area

The various aquifer characteristics within the study area are shown in Table 1. In order to reveal the geologic sections in different parts of the study area, a profile, BB', was selected as shown in figure 2, which traverses Ibeme (KS50), Umueze (KS42), Obohia (KS37) in Ogwashi-Asaba Formation. The section shows the presence of sands and sandstone, figure 3. There is a correlation between the VES result and the lithologs of boreholes drilled at locations KS32 (Avutu Obowo), KS37 (Nkwo Obohia) and KS42(Umueze) in Figure 4. They both show fine-medium and coarse-grained sand. They indicate the presence of productive aquifers in the area. According to the geology of the area, the formation is made up of clays, sands, sandstones. KS37 at Nkwo Obohia features the most productive aquifer system discovered in the study, with saturated thickness of 147m, consisting of thick sandstones. Depth to the water table across the study area was deduced from the VES result. The mean depth to the water table of the aquifers in this zone is 60.57m. The depth to the water table at KS37 is quite

small at 2.5m which obviously confirms the shallow unconfined aquifer system reported for the area in literature (Uma, 1989).

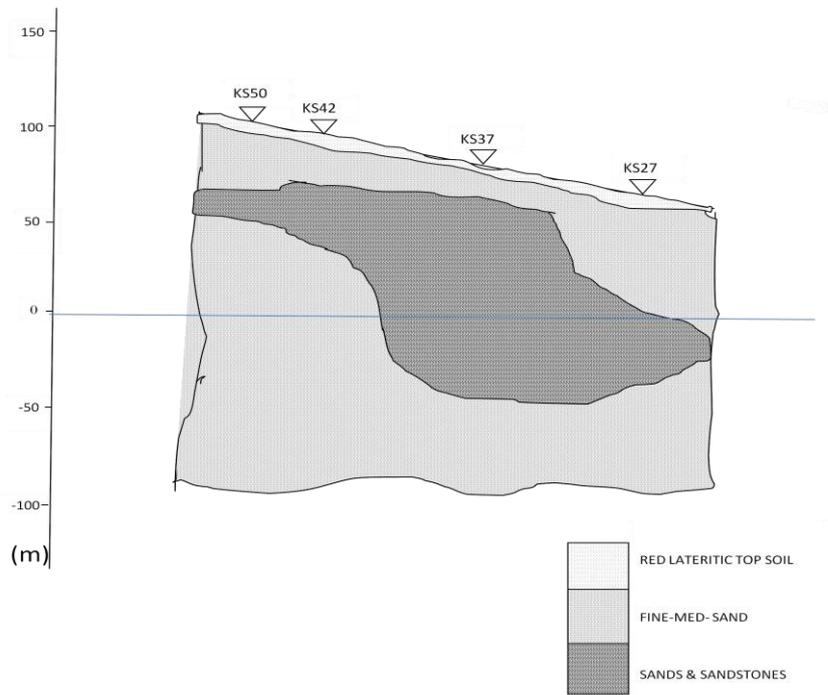


Figure 3: Interpretative Cross-Section along BB'

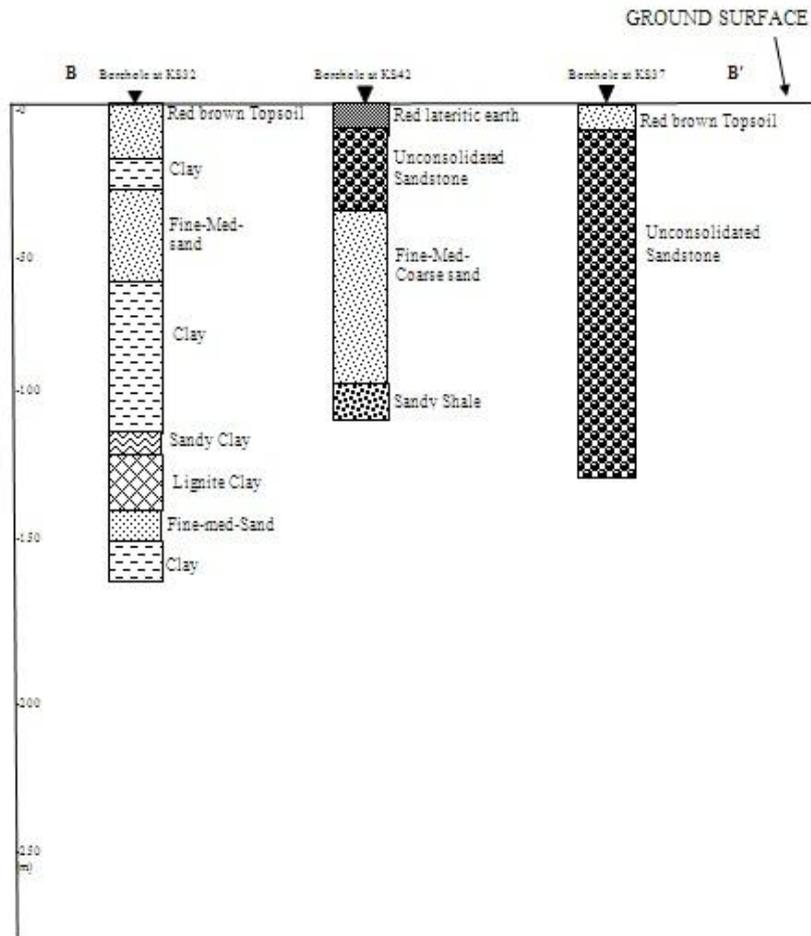


Figure 4: Profile BB' (N-S)

3.2 Hydraulic Characteristics of the Aquifer Systems

The hydraulic characteristics of the aquifer type within the study area were established using the concept of Dar-Zarrouk parameters (transverse unit resistance (R) and longitudinal conductance (C) in porous media. According to Ekwe et al. (2006), Uboma-Obowo-Umuagu is homogeneous hydrologically with $K\sigma$ values varying between 0.0102 and 0.0316, with a mean value of 0.0209 representing the Ogwashi Asaba Formation. Anara-Obohia-Amogwugwu has a mean $K\sigma$ value of 0.0047 representing Ameki Formation, while Imo Formation has a mean $K\sigma$ value of 0.00315. From these values, the hydraulic conductivities of the various VES locations were estimated. From the analysis of Table 1, the following values of hydraulic conductivities and transmissivity were obtained for the aquifers withing Ogwashi-Asaba Formation: $K_{mean}=6.71m/day$; $T_{mean}=327.91m^2/day$; $K_{max}=19.71m/day$; $T_{max}=2906.52m^2/day$; $K_{min}=1.13m/day$; $T_{min}=8.01m^2/day$. The highest values of K and T obtained for the study are for KS37 at Nkwo- Obohia. This indicates that this location has the greatest potential for productive aquifers because of its high transmissivity.

Figure 5 shows the contour map of the hydraulic conductivity obtained for the study area showing areas with high and low hydraulic conductivities. Figure 6 shows the contour map of transmissivity for the study area. These are indicative of the productive potential of the aquifers. Figure 7 shows the contour map of the depth to the water table. Figure 8 shows the contour map of the aquifer thickness. Figure 9 shows the contour map of $K\sigma$ values estimated for the study area.

Table 1: Aquifer Characteristics at the Various VES Locations in the Study Area.

VES #	LOCATION	DEPTH TO WATER TABLE (m)	AQUIFER THICKNESS h (m)	RESISTIVITY $\rho(\Omega\text{-m})$	TRANSVERSE RESISTANCE, R $\rho h (\Omega\text{-m}^2)$	$K\sigma$	HYDRAULIC CONDUCTIVITY $K(m\ day^{-1})$	TRANS-MISSIVITY $K\sigma R (m^2\ day^{-1})$
KS26	AFARA-UKWU, UMUAHIA	36.3	18.2	9541.0	173646	0.0008	7.63	138.92
KS28	UMUOVOR, OLD UMUAHIA	23.8	26.2	7310.0	191522	0.0006	4.39	114.91
KS30	UMUOSCHIE, OBOWO	3.0	10.9	490	5341	0.0062	3.04	33.11
KS32	AVUTU OBOWO	17.4	29.3	5770.0	169061	0.0013	7.50	219.78
KS33	UMUGHARA ONICHA	23.7	43.5	14200	617700	0.0005	7.10	308.85
KS35	NNARAMBIA	13.0	112.0	11254.0	1260448	0.0002	2.25	252.09
KS36	OGBE	13.0	122.0	23568.0	2875296	0.0002	4.71	575.06
KS37	NKWO OBOHIA	2.5	147.5	98526.0	14532585	0.0002	19.71	2906.52
KS38	UMUMBIRI OPARANADIM	40.7	29.0	11400.0	330600	0.0003	3.42	99.18
KS42	UMUEZE	60.2	20.5	4010.0	82205	0.0005	2.01	45.10
KS43	UMUANUNU NSU	37.9	57.1	3211.0	183348	0.0013	4.17	55.00
KS44	UMUAGBA ONICHA	11.6	24	17000.0	408000	0.0008	13.60	326.40
KS48	UMUCHOKO AMACHARA	19.4	100.6	1399	140739	0.0061	8.53	858.51
KS49	ABOH EBIKORO	55.0	30.3	5036	152591	0.0019	9.57	289.92
KS50	NNEATO UMUOKE, IBEME	122.0	57.0	4040.0	230280	0.0006	2.42	138.17
KS53	UMUZOH O EZIHE	30.3	17.8	8300	147740	0.0004	3.32	59.10
KS55	EZIAMA OSUAMA	4.3	2.1	5450.0	11445	0.0007	3.82	8.01
KS60	UMUODIMODU ORODO	13.2	66.8	399.9	26713	0.0061	2.44	162.95
KS62	OKWUDOR	36.2	28.9	10500.0	303450	0.0008	8.40	246.76
KS63	UMUOMA AMUCHA	40.0	62.0	7800.0	483600	0.0016	12.48	773.76
KS64	OFEIYI UMUDI	51.6	34.1	2710.0	92411	0.0017	4.61	157.10
KS65	OBINUHU NKWERRE	23.0	47.5	3210.0	152475	0.0043	13.80	655.64
KS66	UMUGARA NKWERRE	35.5	193.5	188	36378	0.0060	1.13	218.27
KS68	UMUDIWEBU-UMUOPARADIM	65.0	47.0	25000.0	1175000	0.0002	5.00	235.00

KS69	ABBA	21.7	73.3	2800.0	205240	0.0005	1.68	102.62
KS70	UMUANYA- AMAIGBO	14.9	24.4	2400	58560	0.0013	3.12	76.13
KS71	UMULEKE- AMAIGBO	36.6	19.3	10200.0	196860	0.0013	13.26	255.92
KS74	AVUTU-OBOWO	58.0	21.0	1200.0	25200	0.0013	1.56	32.76
KS75	UMUARIAM	42.0	53.0	8608.2	456235	0.0016	13.77	729.96
KS77	AFOR UMUOKEH	59.4	31.6	3570.0	112812	0.0004	1.43	45.12
KS86	UMUDIMOA- AMAIKE	88.4	49.6	6590.0	326864	0.0002	1.32	65.37
KS89	UMUEZE-AMIKE ORLU	41.3	21.1	10800.0	227880	0.0030	32.40	683.60

IV. Conclusion

The diagnostic features of the $K\sigma$ product proved useful in the study. It was used to estimate the hydraulic conductivity and the transmissivity for the sounding locations across the study area. The hydraulic conductivity varies between 1.13m/day (KS66) and 19.71 m/day (KS37). The highest transmissivity value was obtained at Nkwo Obohia in Mbaise area, along profile BB', with 2906.52 m²/day. The aquifer there consists of thick coarse sands and sandstones with a thickness of 122m.

Acknowledgements

The authors wish to acknowledge with gratitude the Anambra-Imo River Basin Authority, the Imo State Rural Water Supply Agency, and UNICEF Owerri for giving us access to their information resources. Special thanks to Engr. Emeka Udokporo of FLAB Engineering and Mr. A. O. Kanu of the Abia State Water Board, who made a lot of useful material available and helped in many other ways; also Mr. R. C. Oty of the Anambra- Imo River Basin Authority, and Mr. R. N. Ibe of the Ministry of Public Utilities, Owerri for their tremendous help in the work.

References

- [1] Dessauvague, T. F. J., (1972); Biostratigraphy of the Odukpani (Cretaceous) Type Section, Nigeria Proc. of the Conf. on African Geology, Ibadan, Nigeria, pp.207-218.
- [2] Ekwe, A. C., Onu N. N., and Onuoha K. M. (2006); Estimation of Aquifer Hydraulic Characteristics from Electrical Sounding Data: The Case of Middle Imo River Basin Aquifers, South-Eastern Nigeria. *Journal of Spatial Hydrology*, Vol. 6, No. 2, pp. 121-132.
- [3] Rubin Y. and Hubbard S.(2005); Hydrogeophysics, Springer, Netherlands, Water and Science Technology Library, V. 50, 523 p., 2005.
- [4] Majumdar R. K. and Das D. (2011); Hydrological Characterization and Estimation of Aquifer Properties from Electrical Sounding Data in Sagar Island Region, South 24 Parganas, West Bengal, India. *Asian Journal of Earth Sciences*, 4(2) pp. 66-74.
- [5] Mendosa, F.G., Steenhuis, S.T., Todd Walter, M., Parlange, J.-Y., (2003). Estimating basin-wide hydraulic parameters of a semi-arid mountainous watershed by recession-flow analysis. *Journal of Hydrology* 279, 57–69.
- [6] Niwas, S., Singhal, D.C., (1981); Estimation of aquifer transmissivity from Dar-Zarrouk parameters in porous media. *Journal of Hydrology* 50, 393–399.
- [7] Niwas, S., Gupta, P.K., de Lima, O.A.L., (2006); Nonlinear electrical response of saturated shaley sand reservoir and its asymptotic approximations. *Geophysics* 71 (3), 129–133.
- [8] Onwuegbuche, A. A., (1993); Geoelectrical investigations in the Imo River Basin Nigeria. Unpublished Ph.D thesis. Department of Physics, University of Calabar. pp.2-62.
- [9] Uma, K.O., (1989); An appraisal of the groundwater resources of the Imo River Basin, Nigeria. *Journal of Mining and Geology* Vol. 25 Nos 1&2, pp.305-31
- [10] Zohdy, A. A. R., Eaton, G.P. and Marbey, D. R., (1974); Application of surface Geophysics to groundwater investigations. US Geological Survey.

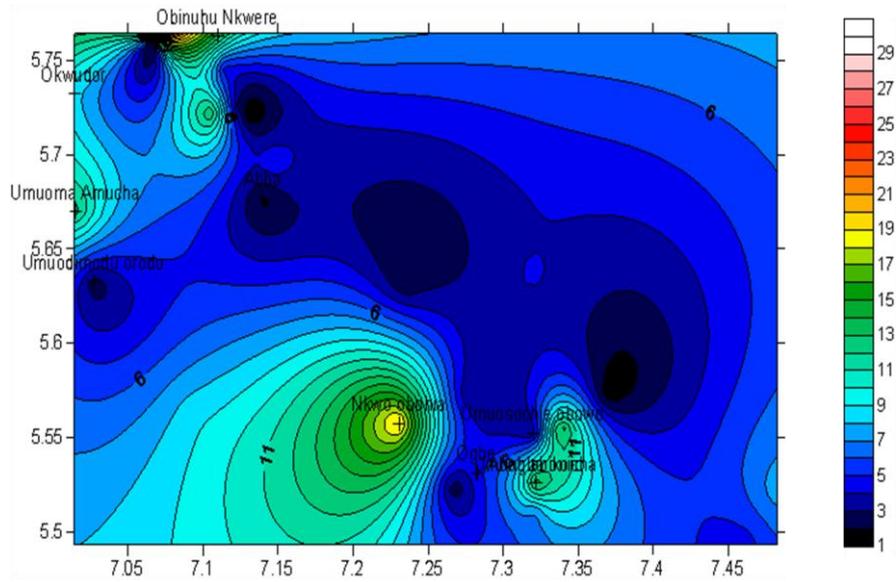


Figure 5: Map of Hydraulic conductivity for the Study Area

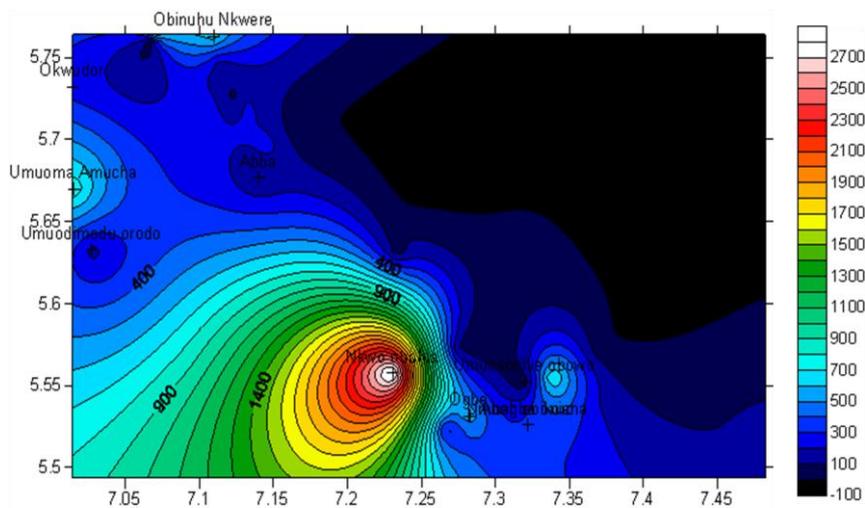


Figure 6: Map of Transmissivity for the Study Area

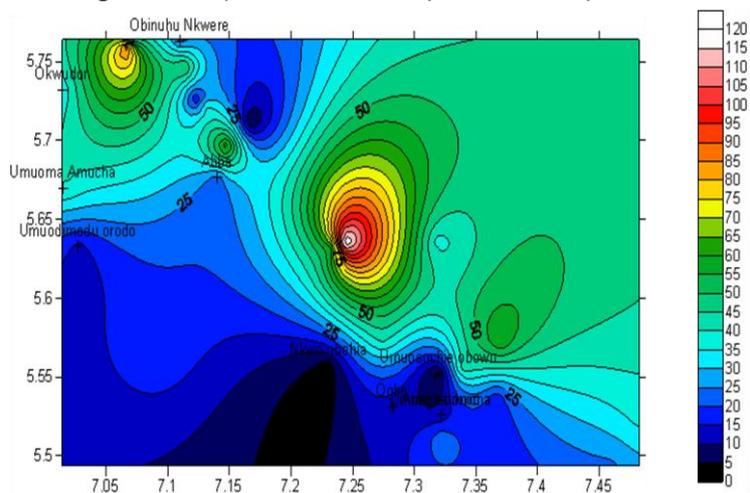


Figure 7: Map showing Depth to Water Table (m)

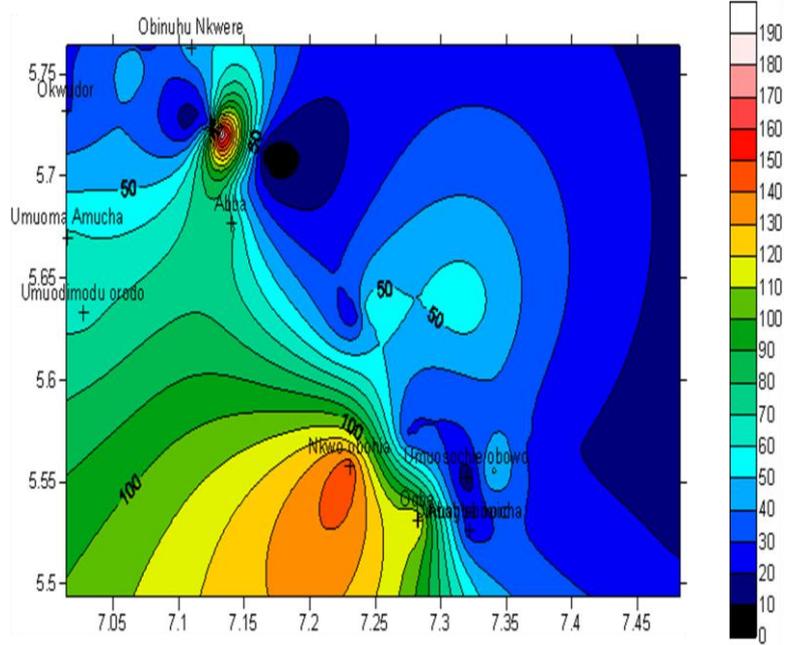


Figure 8: Map of Aquifer thickness(m) for the Study Area

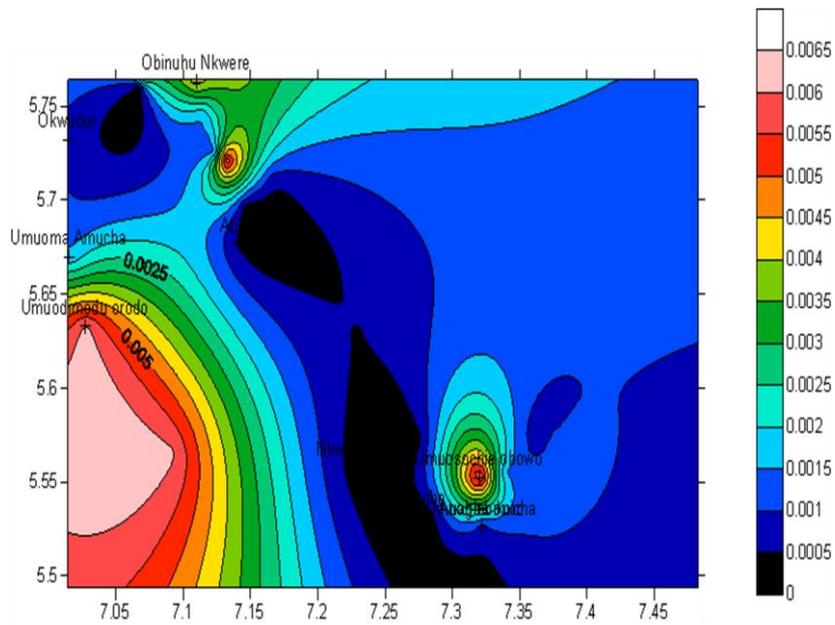


Figure 9: Map of Diagnostic parameter ($K\sigma$)

A Novel Interval Type-2 Fuzzy Software Effort Estimation Using Takagi-Sugeno Fuzzy Controller

V. V. R. Manoj¹, J. N. V. R. Swarup Kumar².

¹ M. Tech (CSE), Gudlavalleru Engineering College, Gudlavalleru.

² Asst. Prof. in Department of CSE, Gudlavalleru Engineering College, Gudlavalleru.

Abstract: Software effort estimation is the process of estimating the cost and time required for the development of software system. Resource allocation and bidding are major parts of planning in software projects. The main objective of the plan is to scout for the future and diagnose the attributes the consummation of successful projects. So, to meet the challenges of cost estimation time in the software development, effective software is required. This paper introduces a novel model of fuzzy logic estimation effort in software development. My paper touches up on MATLAB for tuning parameters of famous various cost estimation models. It also uses published software projects data, performance of the model and the comparison between my novel model and existing ubiquitous models.

Key words: Fuzzy Logic, Effort Estimation, KLOC, COCOMO, Fuzziness, Membership Function.

I. Introduction

Software estimation is the process of predicting the amount of time (effort) that is required to build a software system. The cost benefit analysis is performed with cost estimation process which is achieved in terms as person-months (PM) and can be translated in to actual dollar cost. Estimation carries inherent obscurity risks. The concept of software cost estimation has been growing rapidly these days. Due to globalization, people expect high quality software with a low cost though so many models came into existence like COCOMO81, COCOMOII, SLIM, FP, Delphi, Halsted Equation, Bailey-Basili, Doty, and Anish Mittal Models. Recent surveys repeat says that software projects overrun the cost estimation, which is found in the actual data. COCOMO II is the model is used to estimate the cost as products in many software companies but in vain due to the some variations in models [2], [5-10]. We also can find several problems like unrealistic over-optimum, complexity, and overlooked tasks [11], [12]. So, to overcome all these problems in software development, new models are approached which researchers showed attention in 1990's. They are artificial neural networks, fuzzy logic models and genetic algorithms. Out of these models fuzzy logic is the best powerful linguistic representation with exact inputs and outputs. It is also based on model building with logic concepts introduced by Lofti A. Zadeh [3], [4], [13].

1.1 Membership Functions

Fuzzy numbers are 3 types they are 1) Triangular fuzzy number 2) Trapezoidal fuzzy number 3) Bell shaped fuzzy number. Fuzzy numbers are used to describe the vague data obscurity and imprecision. A fuzzy number is an extension of a regular number that does not refer to one single value

but connected to set of possible values which weights between '0' and '1'. This weight is called the membership function. It increases towards the mean and decreases away from it.

Membership function is characterized by fuzziness in a fuzzy set. It is a curve that defines how each point in the input space is mapped to a membership value between 0 and 1. It may exist different graphical representations but they are certain restrictions regarding shaped to be considered effectively.

Gaussian Bell shape (Figure 1)

It is defined by its mid value m and the value of $k > 0$. The greater k is, the narrower the bell.

$$G(x) = e^{-k(x-m)^2} \dots (1)$$

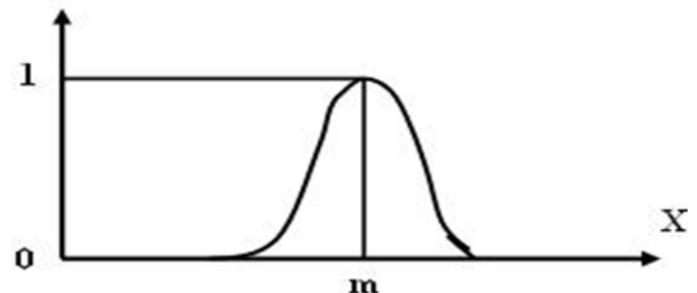


Fig 1: Gaussian Bell Shape

II. COST ESTIMATION MODELS LITERATURE REVIEW

Within last few decades, to improve the accuracy of cost estimation many software cost estimation models [2], [6-10] were introduced. It seems to be impractical because of the inherent obscurity in software development projects and the impact of software development cost use. Still, it is likely that the estimation can be improved because software development cost estimates are systematically overoptimistic and very inconsistent. The primary objective of the software engineers has been to develop required models using which software cost can be accurately estimated. Estimation models use KDLOC (Thousands of Delivered Lines of Code) as the primary input. This input is not sufficient for accurately estimating the cost of products. Several other parameters have to be considered.

2.3 Doty: [1]

$$\text{Effort} = 5.288(\text{KLOC})^{1.047} \dots (2)$$

Halsted Equation:

$$\text{Effort} = 5.2 (\text{KLOC})^{1.50} \dots (3)$$

Bailey-Basili:

$$\text{Effort} = 5.5 + 0.73 (\text{KLOC})^{1.16} \dots (4)$$

2.4 Mittal Model: [1]

2.5 Harish model1

Fuzzification: $u(E) =$

$$\begin{cases} 0 & \text{if } E \leq a \alpha^b \\ \frac{(\frac{E}{a})^{1/b} - \alpha}{m - \alpha} & \text{if } a \alpha^b \leq E \leq am^b \\ \frac{\beta - (\frac{E}{a})^{1/b}}{\beta - m} & \text{if } am^b \leq E \leq a\beta^b \\ 0 & \text{if } E \geq a\beta^b \end{cases} \dots (5)$$

Defuzzification: $E = \frac{(w_1(a\alpha^b) + w_2(am^b) + w_3(a\beta^b))}{(w_1 + w_2 + w_3)} \dots (6)$

2.5 Harish model1

$Effort (E) = \frac{w_1(a\alpha^b) + w_2(am^b) + w_3(a\beta^b)}{w_1 + w_2 + w_3} \dots (7)$

Where $a=3.41, b=0.795, m$ represents size in KLOC

$\alpha = \left(1 - \frac{2KF}{k+1}\right) * m$

$\beta = \left(1 + \frac{2F}{k+1}\right) * m$

K, f, w_1, w_2 and w_3 are arbitrary constants.

The effort is obtained in man months (MM). optimization of effort for an application is done by a suitable choice of arbitrary constants.

Harish model2

$Effort (E) = \frac{w_1(a\alpha^b) + w_2(am^b) + w_3(a\beta^b)}{w_1 + w_2 + w_3} + c(ME) + d \dots (8)$

Where $a=3.41, b=0.795, m$ represents size in KLOC

$\alpha = \left(1 - \frac{2KF}{k+1}\right) * m$

$\beta = \left(1 + \frac{2F}{k+1}\right) * m$

K, f, w_1, w_2, w_3, c and d are arbitrary constants.

ME is methodology of the project.

The effort is obtained in man months (MM). optimization of effort for an application is done by a suitable choice of arbitrary constants.

III. Proposed Model

Interval Type-2 GMF (A,M,B) Firing Intervals:

$J_{Px} = \left[\frac{\mu_{PL1} + \mu_{PL2}}{2}, \frac{\mu_{PR1} + \mu_{PR2}}{2} \right] = [\mu_p(x_i), \mu_p(x_i)] \dots (9)$

$J_{Nx} = \left[\frac{\mu_{NL1} + \mu_{NL2}}{2}, \frac{\mu_{NR1} + \mu_{NR2}}{2} \right] = [\mu_N(x_i), \mu_N(x_i)] \dots (10)$

Defuzzification:

In this model we considered centroid method (weights average), which is of the form [11]

$C = \frac{\sum_{i=1}^N w_i \mu_i}{\sum_{i=1}^N w_i} \dots (11)$

IV. EXPERIMENTAL STUDY:

For the membership functions the L value is the mean of the input sizes i.e. 207.3385 and $stddev()$ is 186.3325 ($L1=393.671, L2= 21.006$). By applying exponential regression [www.xuru.org] analysis for the input sizes and effort we obtained $a=70.737$ and $b=0.004$.

By applying Gaussian membership function for the membership functions the left and right boundaries are $[\mu_p(\alpha, m, \beta), \mu_N(\alpha, m, \beta)]$ measured with $\alpha=0.9m$ and $\beta=1.1m$. Foot print of uncertainty intervals for the μ_p is [0.501212 to 0.612593] for left hand side i.e LMF and [0.9 to 1.1] for right hand side i.e UMF. Foot print of uncertainty intervals for the μ_N is [0.13737 to 0.167896] for left hand side i.e LMF and [0.206148 to 0.251959] for right hand side i.e UMF. The means of FOU intervals is taken as firing strength.

$J_{Px} = [\mu_p(X_i), \mu_p(X_i)] = [0.556903, 1]$

$J_{Nx} = [\mu_N(X_i), \mu_N(X_i)] = [0.152633, 0.229054]$

The type reducer action by using the Gaussian membership function and defuzzification is done through centroid method and results shown in the table II.

V. Research Methodology

The performance of proposed software effort estimation model is evaluated by comparing against various software cost estimation models. The methodology used in empirical evaluation is described as follows:

- For each model, using MRE we evaluate the impact of estimation accuracy using (MARE, VARE) evaluation criteria.
- Criterion for measurement of software effort estimation model performance.

MARE (%) =

$mean \left(\sum_{i=1}^n \left| \frac{estimate_i - actual_i}{actual_i} \right| \right) * 100 \dots (12)$

VARE (%) =

$var \left(\sum_{i=1}^n \left| \frac{estimate_i - actual_i}{actual_i} \right| \right) * 100 \dots (13)$

Where $estimate_i$ is the estimated effort (E) from the model, $actual_i (\hat{E})$ is the actual effort and n is the number of projects.

VI. Model Results & Discussion

S.No	Size	Actual Effort	Doty Model	Balli-Bairath Model	Walton Felix	Harish Model1	Harish Model2	Mittal Model	Harri Model	Swamp Model3
1	39	72	248	86.60212	29.622103	71.792024	84.942024	68.34707	59.666994	77.778217
2	40.5	82.5	254.9	89.992028	20.318064	71.801662	84.653662	70.09411	61.009874	78.070443
3	50	84	317.8	73.753993	24.612704	71.86421	84.61421	80.62635	70.268827	79.366785
4	128.6	230.7	854.4	209.69372	58.144252	72.12628	84.88628	151.11487	121.70216	98.488093
5	163.4	157	1083.8	571.26082	71.497401	72.201861	84.961861	175.7591	153.13061	107.32207
6	164.8	246.9	1107.8	577.76688	72.8667	72.207882	84.967882	178.21264	155.31386	108.97411
7	200	130.3	1365.7	346.31418	86.903113	72.263017	85.013817	202.69681	176.65771	120.70049
8	214.4	86.9	1459.1	374.83974	92.579022	72.283916	85.033916	212.28849	185.01721	125.94762
9	253.6	287	1739.5	454.38993	107.86334	72.332483	85.082483	237.30749	206.87447	141.88708
10	254.2	258.7	1743.9	455.63022	108.09554	72.333167	85.083167	237.74079	207.19983	142.11919
11	289	116	1894.6	577.88332	121.48288	72.370299	85.120299	259.91688	225.62487	159.39888
12	449.9	336.3	3178.3	878.34557	181.72392	72.489537	85.248537	347.25244	302.87943	249.58085
13	450	1107.31	3171.1	878.87062	181.76967	72.489601	85.248601	347.87548	302.9242	249.67659

Table 1: Effort of various models

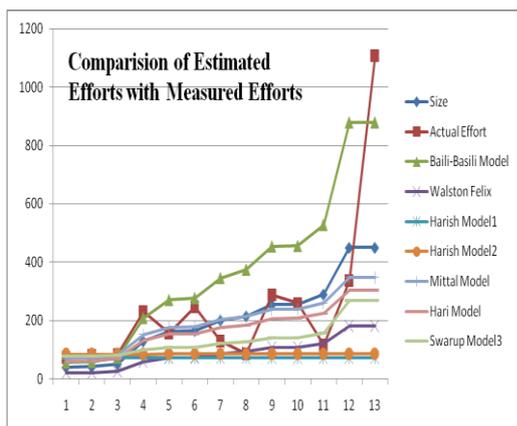


Fig 2: Measured Effort Vs Estimated Efforts of various Models

Comparison of various models on the basis of various performance criterions for software cost estimation is given in below Table 2. Figure 6 below shows the Mean Absolute Relative Error (%) comparison of various models.

Table 2: Various Models MARE % Values

Model Type	MARE %
Doty Model	8186.1
Bailey-Basili	1325
Walston-felix	713.3
Harish Model1	638.8
Harish Model2	563.3
Mittal Model	518.7
Hari Model	515.3
Swarup Model3	442.7

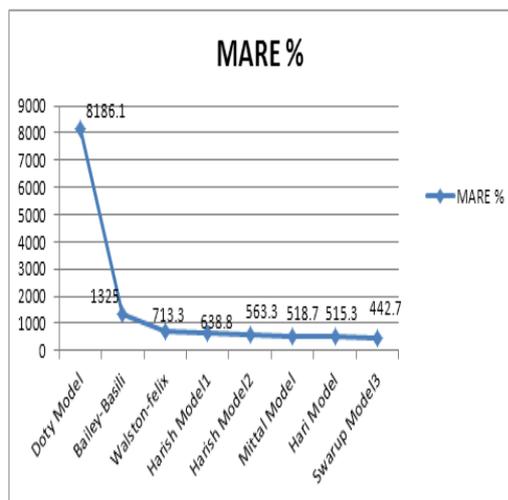


Fig 3: MARE (%) Comparison of various models

VII. Conclusion

Software development life cycle is important for project managers are to estimate the accuracy and reliability at the early stages of software development. This paper postulates about the fuzzy software cost estimation model and with other popular software cost estimation models. It concludes by empirical evaluation of better software effort with proposed and traditional estimation models by MARE

evaluation criteria. To identify the problem of obscurity and vagueness that are existed in software effort drivers' fuzzy logic methods are applied. This proves the fuzzy logic application is used in software engineering successfully.

References

- [1] Anish Mittal, Kamal Parkash, Harish Mittal, "Software Cost Estimation using fuzzy logic", ACM SIGSOFT Software Engineering Notes, November 2010 Volume 35 Number 1.
- [2] Robert W. Zmud, Chris F. Kemerer, "An Empirical Validation of Software Cost Estimation Models" Communication of the ACM Vol 30 No 5, May 1987.
- [3] Zadeh, L.A., Fuzzy sets, Info and Control, 8,338-353, 1965.
- [4] Jose Galindo".Handbook of Research in Fuzzy Information Processing in Databases", Information science Reference, 2008.
- [5] Kim Johnson, Dept of Computer Science, University of Calgary, Alberta, Canada, "Software Cost Estimation: Metrics and Models" (pages 1 to 17).
- [6] CH.V.M.K.Hari et.al, "Identifying the Importance of Software Reuse in COCOMO81, COCOMOII.", International Journal on Computer Science and Engineering Vol.1 (3), 2009, 142-147, ISSN: 0975-3397.
- [7] Baiely,j.w Basili,"A Metamedel for Software Development Resource Expenditure." Proc. Intl. Conference Software Egg. pp : 107-115,1981.
- [8] B.Boehm, Software Engineering Economics Englewood Cliffs, NJ, Prentice Hall, 1981.
- [9] B. Boehm., Cost Models for Future Life Cycle Process: COCOMO2. Annals of Software Engineering. 1995
- [10] Pankaj jalote, "An Integrated Approach for Software Engineering.", Third Edition. ISBN: 978-81-7319-702-4.
- [11] Magne Jorgensen, Stein Grimstad, Simula Research Laboratory, Norway, "Over- Optimism in Software Development Projects: "The Winner's Curse", (CONIELECOMP-2005).
- [12] T. Menzies, D. Port, Z. Chen, J. Hihn, and S. Stukes, "Validation Methods for calibrating software effort models," in ICSE '05: Proceedings of the 27th international conference on Software engineering, (New York, NY, USA), pp. 587-595, ACM Press, 2005.
- [13] Lotfi Zadeh, A., 1994. Fuzzy Logic, Neural Networks and Soft Computing, Communication of ACM., 37(3): 77-84.



V. V. R. MANOJ, He is pursuing his M.tech degree in J.N.T.U. kakinada. He Received B. Tech Degree from ACHARYA NAGARJUNA UNIVERSITY.

Global asymptotic behaviour of aquatic vegetation in a periodical mediated system

Adu A. M. Wasike¹, George E. Kimathi², and Ganesh P. Pokhariyal²

¹ Department of Mathematics, Masinde Muliro University of Science and Technology.

² School of Mathematics, University of Nairobi,

Abstract: In this paper, we consider a competition model between two species in an aquatic system. The operating parameters and the species specific response functions are periodic functions of time. Species-specific removal rates are also permitted to be periodic. A threshold result on the global dynamics of the scalar asymptotically periodic Kolmogorov equation is applied to a growth model of two species. Sufficient conditions ensure uniform persistence of all the species and guarantee that the full system admits at least one positive, periodic solution. The qualitative behaviour of this model is determined analytically and numerically.

Key words: Periodic Kolmogorov equations, asymptotically periodic semiflows.

AMS (MOS) subject classifications: 94A15, 92A17, 34C15, 34C35.

I. Introduction

Mathematical models of a chemostat-like model of two species of vegetation competing exploitatively for an essential, non-reproducing, growth-limiting nutrient whose concentration varies periodically with time predict competitive exclusion. That is, they predict that at most one competitor population avoids extinction (see, e.g., [1], [1], [20], [2]). However, the coexistence of competing populations is common in nature, and so in order to explain this, it seems necessary and natural to introduce periodic coefficients to represent, for example, seasonal variations in the environment.

We consider the chemostat-like model where nutrient supply, specific death rates and species-specific nutrient uptake function are assumed to be periodic with commensurate periods. The model incorporating periodic coefficients takes the following form:

$$\begin{aligned} \dot{S}(t) &= (S^0(t) - S(t))D_0(t) - \sum_{i=1}^2 c_i P_i(t, S(t))x_i(t), \\ \dot{x}_i(t) &= (P_i(t, S(t)) - D_i(t))x_i(t), \\ S(0) &= S_0 > 0, \quad x_i(0) = x_{i0}, \quad i = 1, 2. \end{aligned} \tag{1}$$

where $x_i(t)$ denotes the density of the i^{th} species or biomass at time t , $S(t)$ denotes the nutrient concentration in the water external to the plant cells at time t , $S^0(t)$ is the periodic inflow concentrations at time t , c_i is the content (or concentration) of the nutrient in the plant tissue of species i . Here, $D_0(t)$ is the rate at which the nutrient enters and leaves the aquatic system, and $D_i(t)$ is the specific removal and death rate of the i^{th} species at time t . The function

$$P_i(t, S(t)) := \frac{\mu_i(t)S(t)}{K_i + S(t)},$$

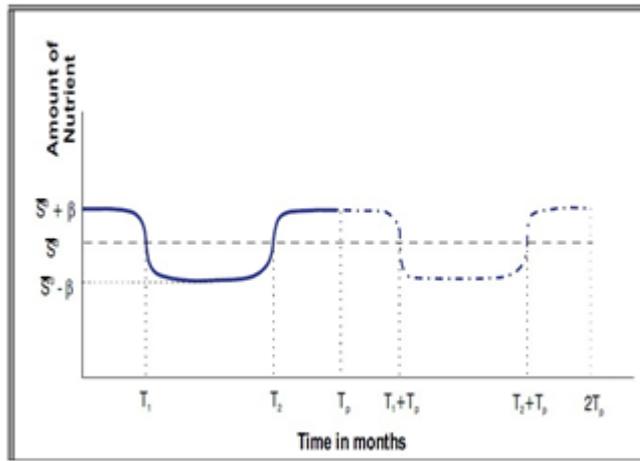
is similar to the Holling type II function [Error! Reference source not found.] and describes the per capita nutrient uptake rate of the i^{th} species, where $\mu_i(t)$ is the maximal specific growth rate of the i^{th} species at time t , K_i is the half saturation

constant for nutrient-limited growth for the i^{th} species. The model structure of Equation 1 that originates from the chemostat theory is suitable for modeling aquatic systems, since all the parameters involved in it can be measured in the field (see [2], [Error! Reference source not found.], [Error! Reference source not found.], [1]).

The periodic coefficients in the model are reasonable. Consider a situation of a fresh water lake that receives its nutrients mainly from streams draining the watershed. As seasons change, stream drainage patterns change causing variations in the supply of nutrients. This is particularly true in many fresh water lakes of East Africa. Moreover, the inflow rate $D_0(t)$ is not

constant but periodic and the period is the same as that of $S^0(t)$. In some cases, the vegetation is fed on by beetles that are their natural biological control. The beetles' rate of reproduction and rate of feeding on the vegetation is positively correlated to the temperature of the season (see [Error! Reference source not found.], [Error! Reference source not found.]). Thus $D_i(t)$, and $\mu_i(t)$, $i = 1, 2$ need to have periods that are commensurate to that of $S^0(t)$. During warmer seasons, the nutrient supply is low while in the colder seasons, the opposite is true. We assume a seasonal input nutrient concentration of the form illustrated in Figure 1.

Figure 1: Variation of Nutrient Concentration with Time



Let $S^0(t)$ represent the periodic input concentration, which fluctuates about a mean value \tilde{S} thus,

$$S^0(t) = \begin{cases} \tilde{S} + \beta, & 0 \leq t \leq T_1; \\ \tilde{S} - \beta, & T_1 \leq t \leq T_2; \\ \tilde{S} + \beta, & T \leq t \leq T_p. \end{cases}$$

The period is T_p , which could be measured in weeks or months and $\beta > 0$ indicates the deviation from the mean value \tilde{S} .

Since $S^0(t)$ sketched in Figure 1 is even, it can be represented by a Fourier series of the form,

$$S^0(t) = \tilde{S} + \beta + 2\beta \frac{(T_1 - T_2)}{T_p} + \alpha(t), \quad (2)$$

where

$$\alpha(t) = \sum_{n=1}^{\infty} \frac{4\beta}{n\pi} \cos\left(\frac{n\pi}{T_p}(T_1 + T_2)\right) \sin\left(\frac{n\pi}{T_p}(T_1 + T_2)\right) \cos\frac{2\pi nt}{T_p}. \quad (3)$$

From Equation (3) we observe that $|\alpha(t)| \leq \frac{4\beta}{\pi}$. The nutrient supplied at any instant satisfies the inequality

$$S^0 - |\alpha(t)| \leq S^0(t) \leq S^0 + |\alpha(t)|$$

We know for practical purposes that $S^0 \leq \frac{4\beta}{\pi}$. This condition means that the amplitude of nutrient fluctuation about \tilde{S} must be small in comparison to S^0 . Also,

$$D_0(t) = \tilde{D}_0 + \beta_0 + 2\beta_0 \frac{(T_1 - T_2)}{T_p} + \alpha_0(t),$$

where \tilde{D}_0 is the mean inflow rate, β_0 indicates the deviation from the mean value \tilde{D}_0 , and $\alpha_0(t)$ is similar to $\alpha(t)$ in (3) with β replaced by β_0 . From the biological and seasonal response thus far discussed, we have that if D_{i0} is the mean value of $D_i(t)$, then

$$D_i(t) = \begin{cases} \tilde{D}_{i0} - \beta_i, & 0 \leq t \leq T_1; \\ \tilde{D}_{i0} + \beta_i, & T_1 \leq t \leq T_2; \\ \tilde{D}_{i0} - \beta_i, & T \leq t \leq T_p, \end{cases}$$

$i = 1, 2$ where $\beta_i > 0$, are non-negative constants and $\alpha_i(t)$ is similar to $\alpha(t)$ in (3) with β replaced by $-\beta_i$. Thus,

$$D_i(t) = \tilde{D}_{i0} + \beta_i + 2\beta_i \frac{(T_1 - T_2)}{T_p} + \alpha_i(t),$$

where, without loss of generality, we have assumed that the period of $D_0(t)$ and $D_i(t)$ are the same. The shape of the $\mu_i(t)$ curve is similar to that of $D_i(t)$.

There has been some research on models similar to that in (1) involving either periodic nutrient input or periodic dilution rates (see [1], [Error! Reference source not found.], [Error! Reference source not found.], [1], [1], [1], [1], [1]). Some discussion on the periodic gradostat has also been considered (see [1], [1]). In most of the previous analytical studies of the periodic chemostat, the powerful theory of monotone dynamical systems was applied to limiting systems obtained using certain conservation principles. However, the theory of monotone dynamical systems can only be applied in this context to study the competition between at most two species. Also, in order to apply a conservation law to obtain the limiting systems, it is necessary to assume that all the removal rates are equal, thus ignoring all the species-specific death rates and only considering the dilution rate.

In this paper, we apply the theory of asymptotically periodic semiflows [Error! Reference source not found.] and the comparison method [Error! Reference source not found.] to determine criteria that guarantees the existence of at least one positive periodic solution for the full system and the uniform persistence of all the interacting species.

This paper is organized as follows. In section 2, we give some preliminaries while in section 3, the two-species model is studied and a statement of the main results is presented. In section 4, we give numerical results that seem to confirm the analytical findings. We conclude with a discussion in section 5.

II. Preliminaries

Consider the n -dimensional Kolmogorov periodic system

$$\dot{u} = uF_0(t, u), \tag{4}$$

Where $u = (u_1, u_2, \dots, u_n) \in \square_+^n$. We assume that $F_0 = (F_{01}, F_{02}, \dots, F_{0n}) : \square_+^{n+1} \rightarrow \square_+^n$ is continuous and ω -periodic with respect to t ($\omega > 0$), and that the solution $\phi_0(t, u)$ of (4) with $\phi_0(0, u) = u$ exists uniquely on $[0, \infty)$. Let

$S = \phi_0(\omega, \cdot) : \square_+^n \rightarrow \square_+^n$. Then $S^m(u) = \phi_0(m\omega, u), \forall u \in \square_+^n$.

Lemma 2.1. *If for some $1 \leq i \leq n$,*

$u^*(t) = (u_1^*(t), \dots, u_{i-1}^*(t), 0, u_{i+1}^*(t), \dots, u_n^*(t))$, *is an ω -periodic solution of (4) with $u_j^*(0) \geq 0, \forall 1 \leq j \leq n, j \neq i$, and*

$u^*(t)$ *satisfies $\int_0^\omega F_{0i}(t, u^*(t)) dt > 0$ then there exists a $\delta > 0$ such that $\limsup_{m \rightarrow \infty} d(S^m(u), u^*(0)) \geq \delta, \forall u \in \text{int}(\square_+^n)$.*

Also consider the n -dimensional non autonomous Kolmogorov system

$$\dot{u} = uF(t, u) \quad 1 \leq i \leq n, \tag{5}$$

Where $u = (u_1, u_2, \dots, u_n) \in \square_+^n$. We assume that $F = (F_1, F_2, \dots, F_n) : \square_+^{n+1} \rightarrow \square_+^n$ is continuous and locally Lipschitz in u . For $s \geq 0$, let $\phi_0(t, s, u)$ and $\phi(s, s, u) = u$ respectively.

Define $T_m := \phi(m\omega, 0, u)$, $T(t)u := \phi_0(t, 0, u)$ and $S(u) := T(\omega)u, \forall m \geq 0, t \geq 0, u \in \square_+^n$.

Lemma 2.2. *Assume that $\lim_{t \rightarrow \infty} |F(t, u) - F_0(t, u)| = 0$ uniformly for u in any bounded subset of \square_+^n , and that solutions of*

(4) and (5) are uniformly bounded in \square_+^n . If for some $1 \leq i \leq n$,

$$u^*(t) = (u_1^*(t), \dots, u_{i-1}^*(t), 0, u_{i+1}^*(t), \dots, u_n^*(t))$$

is an ω -periodic solution of (4) with $u_j^(0) \geq 0, \forall 1 \leq j \leq n, j \neq i$, and $u^*(t)$ satisfies $\int_0^\omega F_{0i}(t, u^*(t)) dt > 0$ then,*

$$\tilde{W}^s(u^*(0)) \cap \text{int}(\square_+^n) = \emptyset,$$

where

$$\tilde{W}^s(u^*(0)) = \left\{ u \in \square_+^n : \lim_{m \rightarrow \infty} T_m(u) = u^*(0) \right\}.$$

Consider equations (4) and (5) with $n = 1$. This will represent a single population growth model. Assume that

(A1) $\lim_{t \rightarrow \infty} |F(t, u) - F_0(t, u)| = 0$ uniformly for u in any bounded subset of \square_+^n , and

there exists $K > 0$ such that $F(t, u) \leq 0, t \geq 0, u \geq K$;

(A2) For any $t \geq 0, F_0(t, u)$ is strictly decreasing for u , and there exists $K_0 > 0$

such that $F_0(t, K_0), t \geq 0$.

We have the following threshold dynamics for the asymptotically periodic equation (5) with $n = 1$.

Lemma 2.3. Assume that A1 and A2 hold

(a) If $\int_0^\omega F_0(t, 0)dt \leq 0$, then $\lim_{t \rightarrow \infty} \phi(t, 0, u) = 0, \forall u \in \mathbb{R}_+$;

(b) $\int_0^\omega F_0(t, 0)dt > 0$ then $\lim_{t \rightarrow \infty} (\phi(t, 0, u) - u^*(t)) = 0, \forall u \in \mathbb{R}_+ \setminus \{0\}$, where $u^*(t)$ is the unique positive ω -periodic solution of the periodic Kolmogorov equation (4) with $n = 1$.

The reader is referred to [1] for proof of these lemmas.

III. Analysis of the Model

If we let $\bar{x}_i(t) = c_i x_i(t)$ and ignore the bars for notational brevity, system (1) becomes

$$\begin{aligned} \dot{S}(t) &= (S^0(t) - S(t))D_0(t) - \sum_{i=1}^2 P_i(t, S(t))x_i(t), \\ \dot{x}_i(t) &= (P_i(t, S(t)) - D_i(t))x_i(t), \\ S(0) &= S_0 > 0, x_i(0) = x_{i0}, i = 1, 2. \end{aligned} \tag{6}$$

Throughout this paper, we identify the unique solution of Equation (6) by the set

$$(S(t), x_1(t), x_2(t)) \in \mathbb{R}_+^3, S \geq 0, x_i \geq 0, 1 \leq i \leq 2,$$

Where \mathbb{R}_+^3 is a real 3-dimensional non-negative vector space. By asserting that $(S(t), x_1(t), x_2(t))$ is positive, we mean that each component of the solution is positive for all $t > 0$.

The following results found, for example, in [1] will be useful in our analysis.

Let $D(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a continuous, ω -periodic, and positive function. The linear periodic equation

$$\dot{V}(t) = S^0(t)D_0(t) - D(t)V(t), V(0) \geq 0, \tag{7}$$

has a unique positive ω -periodic solution $V^*(t)$ such that every solution $V(t)$ of (7) with $V(0) \geq 0$ satisfies

$$\lim_{t \rightarrow \infty} (V(t) - V^*(t)) = 0.$$

Indeed, $V^*(t)$ is given by

$$V^*(t) = e^{-\int_0^t D(s)ds} \left[\frac{\int_0^\omega e^{\int_0^s D(u)du} S^0(s)D_0(s)ds}{\int_0^\omega D(s)ds - 1} + \int_0^t e^{\int_0^s D(u)du} S^0(s)D_0(s)ds \right].$$

Let $\bar{D}(t) = \max\{D_0(t), D_1(t), D_2(t)\}$ and $\underline{D}(t) = \min\{D_0(t), D_1(t), D_2(t)\}$. Clearly $\bar{D}(t)$ and $\underline{D}(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$

are continuous, ω -periodic and positive functions. Let $V_1^*(t)$ and $V_2^*(t)$ be the unique positive ω -periodic solutions of (7)

with $D(t)$ replaced by $\underline{D}(t)$ and $\bar{D}(t)$ respectively. By the comparison theorem and the global attractivity of

$V_i^*(t), i = 1, 2$, it easily follows that $V_2^*(t) \leq V_1^*(t), \forall t \geq 0$.

The following result concerns the periodic solution of the model in Equation (6).

Theorem 3.1. Assume that

1) $\int_0^\omega (P_i(t, V_1^*(t)) - D_i(t))dt > 0, i = 1, 2;$

2) $\int_0^\omega (P_i(t, V_2^*(t)) - x_j(t)) - D_i(t)dt > 0, i, j = 1, 2, i \neq j$ where $x_j^*(t)$ is the unique positive, ω -periodic

solution of the scalar periodic equation $\dot{x}_j = x_j(P_j(t, V_1^*(t)) - x_j) - D_j(t), 1 \leq j \leq n$. Then system (6) admits a

positive ω -periodic solution, and there exists $\alpha > 0$ and $\beta > 0$ such that any solution $(S(t), x_1(t), x_2(t))$ of (6) satisfies

$$0 < \alpha \leq \liminf_{t \rightarrow \infty} x_i(t) \leq \limsup_{t \rightarrow \infty} x_i(t) \leq \beta, i = 1, 2.$$

Proof. We first show that the solution to (6) is positive.

By Lemma 2.3, condition (1) implies that the periodic equation

$$\dot{x}_i(P_i(t, V_1^*(t)) - x_i) - D_i(t), i = 1, 2,$$

admits a unique ω -periodic solution $x_i(t)$ that is globally attractive in $\square_+ \setminus \{0\}$. We further claim that

$V_1^*(t) > x_i^*(t), \forall t \in [0, \omega]$. Indeed, let $x_i^*(t_1) = \max_{0 \leq t \leq \omega} x_i^*(t), t_1 \in [0, \omega], i = 1, 2$. Then $\dot{x}_i^*(t_1) = 0$, and hence

$$P(t_1, V_1^*(t_1) - x_i^*(t_1)) = D_1(t_1) > 0.$$

Since $P(t, s)$ is strictly increasing for $s \in \square_+$, $V_1^*(t_1) > x_i^*(t_1)$. Let $y(t) = V_1^*(t) - x_1^*(t)$. Then $y(t)$ satisfies the periodic differential equation

$$\dot{y} = S^0(t)D_0(t) - \underline{D}(t)V_1^*(t) - (V_1^*(t) - y)(P(t, y) - D_1(t)). \quad (8)$$

Since $y(t_1) > 0$ and

$$\dot{y}|_{y=0} = S^0(t)D_0(t) + (D_1(t) - \underline{D}(t))V_1^*(t) \geq S^0(t)D_0(t) > 0$$

it follows that $y(t) > 0, \forall t \geq t_1$. Thus the ω -periodicity of $y(t)$ implies that $y(t) > 0, \forall t \geq 0$; that is,

$$V_1^*(t) > x_1^*(t), \forall t \geq 0.$$

For any $(S_0, x_0) = (S(0), x_1(0), x_2(0)) \in \square_+^3$ with $x_i(0) > 0, i = 1, 2$, let $(S(t), x(t)) = (S(t), x_1(t), x_2(t))$ be the unique solution of (6) on the maximal interval of existence $[0, t_+)$. Since $\dot{S}(t)|_{s=0} = S^0(t)D_0(t) > 0$, it follows that $S(t) > 0$, and $x(t) > 0, \forall t \in [0, t_+)$.

We now show that the solution is bounded.

Let $V(t) := S(t) + x_1(t) + x_2(t)$. Then

$$S^0(t)D_0(t) - \bar{D}(t)V(t) \leq \dot{V}(t) \leq S^0(t)D_0(t) - \underline{D}(t)V(t)$$

Therefore, by the comparison theorem, we get

$$\underline{V}(t) \leq V(t) \leq \bar{V}(t), \forall t \in [0, t_+), \quad (9)$$

where $\bar{V}(t)$ is the unique solution of the linear ω -periodic equation

$$\dot{\bar{V}}(t) = S^0(t)D_0(t) - \underline{D}(t)\bar{V}(t)$$

with $\bar{V}(0) = V(0)$, and $\underline{V}(t)$ is the unique solution of the linear ω -periodic equation

$$\dot{\underline{V}}(t) = S^0(t)D_0(t) - \bar{D}(t)\underline{V}(t)$$

with $\underline{V}(0) = V(0)$. The global existence of $\bar{V}(t)$ on $[0, \infty)$ implies that $t_+ = \infty$. Since $\lim_{t \rightarrow \infty} (\bar{V}(t) - V_1^*(t)) = 0$, $V(t)$ and

hence $S(t)$ and $x(t)$ are ultimately bounded; that is, system (6) is point dissipative on \square_+^3 . Therefore, for all $t \geq 0, i = 1, 2$,

$$\dot{x}_i(t) = x_i(t) \left(P_i(t, V(t) - \sum_{j=1}^2 x_j(t)) - x_i(t) - D_i(t) \right) \leq x_i(t) (P_i(t, V(t)) - D_i(t)).$$

By the comparison theorem, it follows that

$$x_i(t) \leq \bar{x}_i(t), \forall t \geq 0, i = 1, 2, \quad (10)$$

where $\bar{x}_i(t)$ is the unique solution of the non autonomous equation

$$\dot{\bar{x}}_i(t) = x_i(t) (P_i(t, \bar{V}(t) - x_i(t)) - D_i(t)), \quad \bar{x}_i(0) = x_i(0) > 0, i = 1, 2. \quad (11)$$

Since $\lim_{t \rightarrow \infty} (\bar{V}(t) - V_1^*(t)) = 0$, we get

$$\lim_{t \rightarrow \infty} (P_i(t, \bar{V}(t) - x_i(t)) - (P_i(t, V_1^*(t)) - x_i(t))) = 0$$

uniformly for x_i in any bounded subset of \square_+ . Since

$$\int_0^\omega (P_i(t, V_1^*(t)) - D_i(t)) dt > 0,$$

by lemma 2.3b we have that

$$\lim_{t \rightarrow \infty} (\bar{x}_i(t) - x_i^*(t)) = 0, i = 1, 2 \quad (12)$$

By (9) and (10), it then follows that for any $i = 1, 2$, and $t \geq 0$,

$$\dot{x}_i(t) = x_i(t) \left(P_i(t, V(t) - \sum_{j=1}^2 x_j(t)) - D_i(t) \right) \geq x_i(t) (P_i(t, \underline{V}(t)) - \bar{x}_j(t) - x_i(t) - D_i(t)), \quad (13)$$

$i, j = 1, 2, \quad i \neq j$

and hence, by the comparison theorem,

$$x_i(t) \geq \underline{x}_i(t), \quad t \geq 0, \quad i = 1, 2, \quad (14)$$

where $\underline{x}_i(t)$ is the unique solution of the non autonomous equation

$$\dot{x}_i(t) = x_i(t) (P_i(t, \underline{V}(t) - x_j(t) - x_i(t)) - D_i(t)), \quad i, j = 1, 2 \quad i \neq j \quad (15)$$

with $\underline{x}_i(0) = x_i(0) > 0, \quad i = 1, 2$. Since $\lim_{t \rightarrow \infty} (\underline{V}(t) - V_2^*(t)) = 0$, we have

$$\lim_{t \rightarrow \infty} (P_i(t, \underline{V}(t) - x_j^*(t) - x_i(t)) - P_i(t, V_2^*(t) - x_j^*(t) - x_i(t))) = 0, \quad i, j = 1, 2 \quad i \neq j \quad (16)$$

uniformly for x_i in any bounded subset of \square_+ . Since

$$\int_0^\omega (P_i(t, V_2^*(t) - x_j(t)) - D_i(t)) dt > 0, \quad i, j = 1, 2, \quad i \neq j$$

by lemma 2.3b we have that

$$\lim_{t \rightarrow \infty} (\underline{x}_i(t) - \underline{x}_i^*(t)) = 0, \quad i = 1, 2 \quad (17)$$

where $\underline{x}_i^*(t), \quad i = 1, 2$, is the unique positive ω -periodic solution of the periodic equation

$$\dot{x}_i(t) = x_i(t) (P_i(t, V_2^*(t) - x_j^*(t) - x_i(t)) - D_i(t)), \quad i, j = 1, 2, \quad i \neq j \quad (18)$$

By (10), (12), (14), and (17), it then follows that

$$\liminf_{t \rightarrow \infty} (x_i(t) - \underline{x}_i^*(t)) \geq 0 \geq \limsup_{t \rightarrow \infty} (x_i(t) - x_i^*(t)), \quad i = 1, 2. \quad (19)$$

Clearly, (19) implies that there exists $\alpha > 0$ and $\beta > 0$ such that any solution $(S(t), x(t))$ of (6) with $S(0) \geq 0$ and $x_i(0) > 0, \quad i = 1, 2$ satisfies

$$0 < \alpha \leq \liminf_{t \rightarrow \infty} x_i(t) \leq \limsup_{t \rightarrow \infty} x_i(t) \leq \beta, \quad i = 1, 2.$$

We now prove the existence of a positive ω -periodic solution to equation (6).

Let $X := \square_+^3$,

$$X_0 := \{(S(t), x_1(t), x_2(t)) \in \square_+^3 : x_i > 0, \quad \forall 1 \leq i \leq 2\},$$

and

$$\partial X_0 := \{(S(t), x_1(t), x_2(t)) \in \square_+^3 : x_i > 0, \quad \text{for some } 1 \leq i \leq 2\}.$$

Then, $X = X_0 \cup \partial X_0$. For any $y = (S_0, x_0) \in X$, equation (6) has a unique solution, $\phi(t, y)$, with $\phi(0, y) = y$.

The map $T(t) = \phi(t, \cdot) : X \rightarrow X$ is a periodic semiflow [2] and $T(t)X_0 \subset X_0, \quad \forall t \geq 0$. We also know that $T(t)$ is point dissipative; that is ultimately bounded in X and uniformly persistent with respect to $(X_0, \partial X_0)$, in the sense that there exists $\eta > 0$ such that for any $y \in X_0, \quad \liminf_{t \rightarrow \infty} d(T(t)y, \partial X_0) \geq \eta$. Let $Q = T(\omega) : X \rightarrow X$ be the Poincare map associated with equation (6). Then by [22] theorem 8.5, the ultimate boundedness of solutions of a periodic system of ordinary differential equations implies the uniform boundedness of solutions, and hence $Q : X \rightarrow X$ is compact.

Therefore, by [2] theorem 2.3, Q admits a fixed point $y_0 \in X_0$ and hence $\phi(t, y_0)$ is a periodic solution of equation (6). Let

$y_0 = (S(0), x_1(0), x_2(0)) \in X_0$. Then $S_0 > 0; \quad x_1(0) > 0$ and $x_2(0) > 0$. It then follows that

$\phi(t, y_0) = (S(t), x_1(t), x_2(t))$ satisfies $S(t) > 0$ and $x_1(t) > 0, \quad x_2(t) > 0$. Consequently, $\phi(t, y_0)$ is a positive ω -periodic solution of (6).

4 Numerical Results

In our model, the nutrient supply and specific death rates are assumed to be periodic with commensurate periods. The nutrient input concentration is described by (2) and (3). We use the following explicit Fourier series to describe the nutrient input function:-

$$S^0(t) = \tilde{S}^0 + \beta \sum_{j=1}^n \cos(j\omega t) \quad (20)$$

As an example, if we set $\tilde{S}^0 = 11$, $n = 9$ we obtain Figure 2 that describes the input nutrient concentrations:-

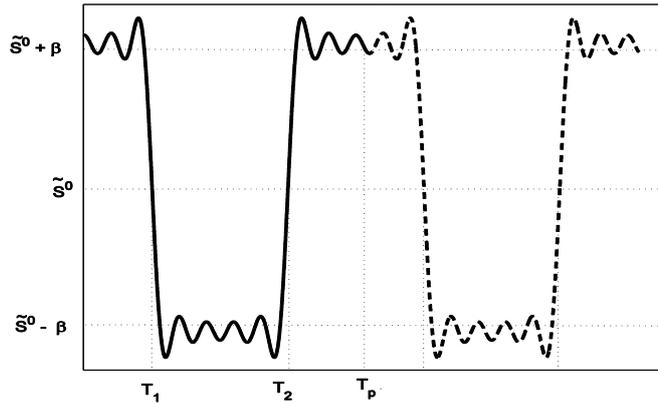


Figure 2: Nutrient Input as described by (20)

Clearly, this figure approximates that envisaged by Figure 1. The level of accuracy of the function may be increased by adding more harmonics to the function, that is making n to be as large as we desire.

It is reasonable to take input nutrient supply of the form

$$D^0(t) = \tilde{D}_0 - \alpha \sum_{j=1}^n \cos(j\omega t), \tag{21}$$

so that the nutrient supply and death rates have commensurate periods. For practical purposes, the fluctuations of the nutrient supply must not be very large, and specifically, we choose our values such that (20) is non-negative, that is

$$\tilde{S}^0 \geq \left| \beta \sum_{j=1}^n \cos(j\omega t) \right|. \text{ For simplicity, we let } D_0(t) = D_1(t) = D_2(t) = \tilde{D}_0 - \alpha \sum_{j=1}^n \cos(j\omega t).$$

We have assumed a Holling type II function for species specific nutrient uptake function $P_i(t, S(t))$ defined by

$$P_i(t, S(t)) = \frac{\mu_i(t)S(t)}{K_i + S(t)}. \text{ With these, (6) is explicitly given by}$$

$$\dot{S}(t) = \left(\tilde{S}^0 + \beta \sum_{j=1}^n \cos(j\omega t) - S(t) \right) \left(\tilde{D}_0 - \alpha \sum_{j=1}^n \cos(j\omega t) \right) - \frac{x_1 \mu_1(t) S(t)}{K_1 + S(t)} - \frac{x_2 \mu_2(t) S(t)}{K_2 + S(t)}$$

$$\dot{x}_1(t) = \left(\frac{\mu_1(t)S(t)}{K_1 + S(t)} - (\tilde{D}_0 - \alpha \sum_{j=1}^n \cos(j\omega t)) \right) x_1(t) \tag{22}$$

$$\dot{x}_2(t) = \left(\frac{\mu_2(t)S(t)}{K_2 + S(t)} - (\tilde{D}_0 - \alpha \sum_{j=1}^n \cos(j\omega t)) \right) x_2(t).$$

We then select the following parameter values that satisfy Lemma 2.3b that give the time plot of numerical solutions of (22) given by Figure 3.

Table 1. Parameter values used in (22) for Figure 3

Parameter	D_0	β	S^0	ω	μ_1	K_1	μ_2	K_2	$x_1(0)$	$x_2(0)$	n	α
Value	0.4675	$\frac{4}{\pi}$	11	1.5	1	1	0.7	0.4	10	10	9	0.3

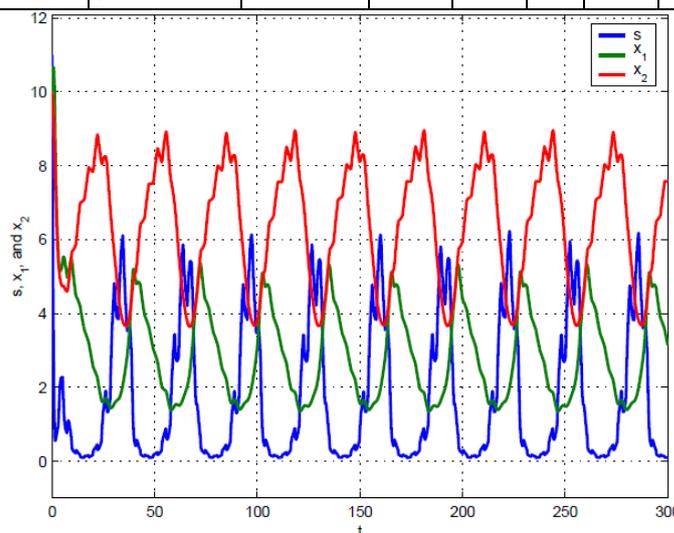


Figure 3: Time Plot for (22)

Figure 3 shows a more realistic existence of each of the competing species as well as the nutrient. As is observed in nature, there are some perturbations in the species at all times. This is reasonable given that deaths occur at all times even when the population is increasing. It is also naturally observed that there will be small increases in population when the general trend is a decline in species biomass.

The numerical results above are confirmed by the following 3-D plot of $S(t)$, $x_1(t)$, and $x_2(t)$.

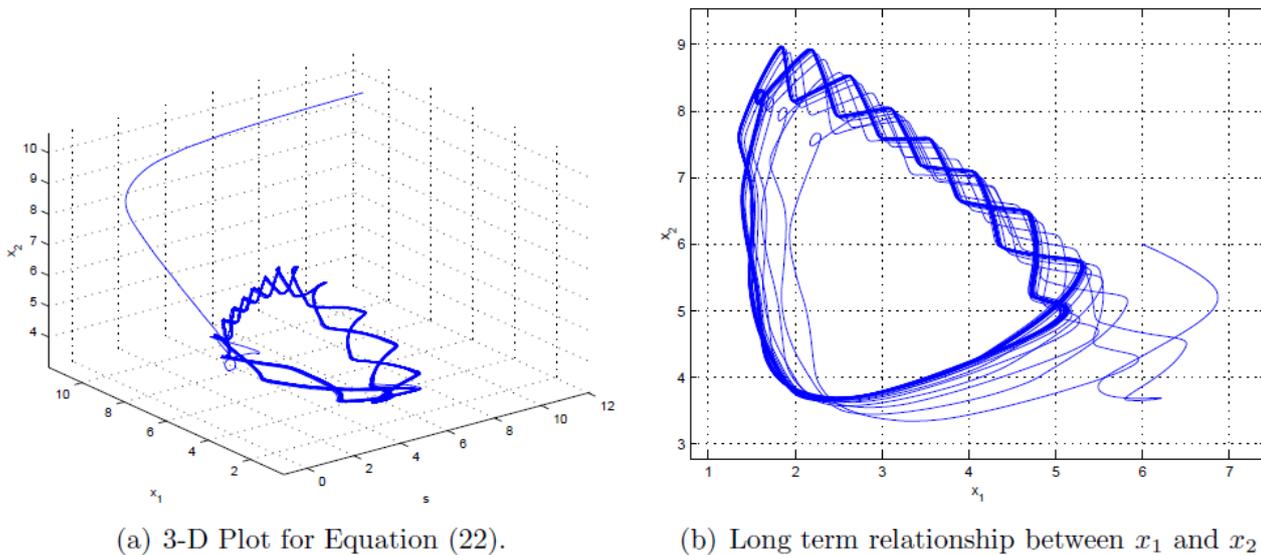


Figure 4: Biomass of x_1 and x_2

We can clearly see from Figures 3 and 4(a) that the solution of (22) remains in the positive interior of (S, x_1, x_2) , meaning that $(S(t), x_1(t), x_2(t)) \in \text{int } \mathbb{R}_+^3$, $0 \leq t \leq \infty$. This behaviour is consistent with what we have determined analytically. The fact that both competing species do not explode (grow without bound) and do not become extinct is clearly demonstrated by the 2-D time plot in Figure 4(b) of x_1 against x_2 . These results are important because models with non-periodic coefficients consistently predict competitive exclusion while we observe that in nature, species competing for a nutrient often exist.

IV. Conclusion

In this paper, we have shown, both analytically and numerically, that (6) admits at least one positive, periodic solution that ensures uniform persistence of all the competing species as predicted by Theorem 3.1. The choice of Fourier series for input and dilution rates is reasonable. In nature, during the warm season, there are temperature variations during the day and at night. The same applies during the cold season. For the case of fresh water lake, assuming nutrients are delivered by run off from rain water, we expect more nutrients during the wet season. However, during this season, it does not rain everyday. Hence we expect some variations in input concentrations within this time as well. The same case would apply during the dry season. During this time, there are usually some light drizzles that would keep the input nutrient concentration fluctuating. In addition, we can modify various parameters in the Fourier series to model specific cases by increasing the amplitude or the period of the function.

Models of the chemostat that assume constant parameters always predicts competitive exclusion, [2]. Our model predicts co-existence and agrees more with what is observed in nature.

Acknowledgments

The authors wish to thank African Millennium Mathematics Science Initiative (AMMSI) and the National Council of Science and Technology (grant No NCST/5/003/PG/171) for their kind support to this research.

References

- [1] G. J. Butler, S. B. Hsu, and P. Watman, *A Mathematical Model of the Chemostat with Periodic Washout rate*, SIAM J. Applied Math, 45 (1985), 435 - 449.
- [2] W. A. Coppel, *Stability and Asymptotic Behaviour of Differential Equations*, D. C. Heath and Co., (1965), Eaglewood.
- [3] C. J. Deloach and H. A. Cordo, *Life Cycle and Biology of Neochentina bruchi and N. Eichorniae on Waterhyacinth in Argentina*. J. of Aquatic Plant Management, 69 (1976)4, 643-652.

- [4] C. J. Deloach and H. A. Cordo, *Life Cycle and Biology of Neochentina Bruchi, a Weevil Attacking Waterhyacinth in Argentina with Notes on N. Eichhorniae*. Annals of the Entomological Society of America, 69 (1976) 4, 643 - 652.
- [5] P. deMottoni and A. Schiafino, *Competition Systems with Periodic Coefficients: A Geometrical Approach*. J. Math. Biology, 11 (1982), 319 - 335.
- [6] S. B. Hsu, S. Hubbell and P. Watman, *A mathematical Theory for Single Nutrient Competition in Continuous Cultures of Microorganisms*. SIAM J. Applied Math, 32 (1977), 366-383.
- [7]
- [8] J. Monod, *Recherches Sur la Croissance des Bacterienness*. Herman, 1942.
- [9] J. K. Hale and A. S. Somolinas, *Competition for Fluctuating Nutrient* J. Math. Biology, 18 (1983), 255 - 280.
- [10] S. B. Hsu, *A Competition Model for a seasonally Fluctuating Nutrient*. J. Math. Biology, 9 (1981), 115 - 132.
- [11] H. L. Smith, *Competitive Coexistence in an Oscillating Chemostat*. SIAM J. Applied Mathematics, 40 (1981), 498 - 522
- [12] H. L. Smith, *Microbial Growth in Periodic Gradostats*. Rocky Mountain J. Math., 20 (1990), 1173 - 1194.3
- [13] H. L. Smith, *Periodic Rotating Waves in a Model of Microbial Competition in a Circular Gradostat*. Canadian Applied Math. Quarterly, 1(1993), 83-113.
- [14] J. W. Hso and P. Waltman, *A Non-linear Boundary Value Problem Arising From Competition in the Chemostat*. Applied Maths and Computation, 32 (1989), 169 - 183.
- [15] H. L. Smith and P. Waltman, *The Theory of the Chemostat*. Cambridge University Press, 1995.
- [16] G. Stephanopoulos, A. G. Fredrickson and R. Aris, *The Growth of Competing Microbial Populations in CSTR with Periodically Varying Inputs*. American Institute of Chemical Engineering J., 25(1979), 863 - 872.
- [17] F. Yang and H. I. Freedman, *Competing Predators for a Prey in a Chemostat Model with Periodic Nutrient Input*. J. Mathematical Biology, 29 (1991), 715 - 732.
- [18] A. Wasike, N. Muiruri and G. Pokhariyal, *Global Assymptotic Behavior of an Aquatic System with Periodic Nutrient Supply*. AJST, 7 (2006) 4, 41 - 50.
- [19] F. J. Weissing, *Biodiversity of Plankton by Species Oscillations ad Chaos*. Nature, 402(1999), 407-410
- [20] G. Wolkowicz and Z. Lu, *Global Dynamics of a Mathematical Model of Competition in the Chemostat: General Response Functions and Differential Death Rates*. SIAM J. Applied Mathematics. 52 (1992), 222 - 233.
- [21] G. S. K. Wolkowicz and H. Xia, *Global Asymptotic behavior of a Chemostat Model with Discrete Delays*, SIAM J. Applied Math, 57 (1997) 4, 1019 - 1043.
- [22] T. Yoshizawa, *Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions*. Applied Mathematical Sciences 14, Springer-Verlag, 1975.
- [23] X. Q. Zhao, *Asymptotic Behavior for Asymptotically Periodic Semiflows with Applications*. Communications on Applied Non-Linear Analysis, 3 (1996) 4, 43 - 66.
- [24] X. Q. Zhao, *Uniform Persistence and Periodic Co-existence States in Infinite-dimensional Periodic Semiflows with Applications*. Canadian Applied Mathematics Quarterly, 3 (1995), 473 - 495.

Investigation of Crystallization Kinetic Parameters of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ Glass Ceramic in presence of Various Nuclei

¹Narmin Montakhab, ²Parisa Sowti Khiabani, ³Mohammad Rezvani

¹Department of Materials Engineering, Faculty of Mechanical Engineering, University of Tabriz, Tabriz, Iran

ABSTRACT: The effect of Y_2O_3 , CeO_2 , P_2O_5 , ZrO_2 and TiO_2 in single, double, triple and quadruple forms on crystallization mechanism of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ (LAS) glass-ceramic system is investigated in this article. Nucleation and crystallization temperatures of optimized samples in each form were determined by Ray & Day method. The crystalline phase and microstructures of samples were studied by X-ray diffraction (XRD) analysis and Scanning Electron Microscopy (SEM) images, respectively. Activation energy of crystallization, E , Avrami and kinetic constants (n , m) were determined by using Differential Thermal Analysis (DTA) and Marotta and Augis-Bennet methods. According to the obtained kinetic parameters (m, n), glasses contained both ZrO_2 and TiO_2 nuclei showed bulk crystallization. The samples contained ZrO_2 , TiO_2 and CeO_2 in the triple nuclei forms, showed two-dimensional bulk crystallization mechanism. Lowest activation energy of crystallization, E (255.5KJ/mol), and highest Avrami constant, n (4.38), were obtained from the sample contained 3wt% TiO_2 and 1wt% ZrO_2 . Lattice constants of the main phase (β -eucryptite solid solution) in the obtained samples were determined according to their XRD results.

Keywords: LAS glass ceramics; Eucryptite; h-quartz; Kinetic parameters; nucleation and crystallization mechanism

I. INTRODUCTION

Glass-ceramic materials are mainly produced via two controlled stages of nucleation and crystallization. These materials have some advantages, such as minimal or even zero porosity and homogeneous microstructure, over the ceramics that are produced via powder metallurgy [1]. These advantageous properties are achieved by precipitating a large percentage of desired crystal phases (at least 50% vol.) in glass-ceramics[2].

Lithium aluminum silicate (LAS) glass-ceramics are one of the most important glass-ceramic systems which have been extensively investigated and commercialized because of their low (zero or negative) thermal expansion coefficient as well as excellent thermal and chemical durability [3-10].

The most important stable crystalline phases in LAS glass-ceramic systems are Eucryptite, Spodumene, Petalite and meta-stable solid solutions such as β -Eucryptite (h-quartz) and Keatite (tetragonal SiO_2) [1]. Investigations have shown that the most effective nucleating agents in these systems are TiO_2 , ZrO_2 , Fe_2O_3 , Cr_2O_3 , NiO , ZnO , V_2O_5 , P_2O_5 and Ta_2O_5 [11].

Sung et al [12] have introduced TiO_2 as an effective nucleating agent in $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glass-ceramic systems. According to their results, the crystallization activation energy of a sample containing 3.85 wt% TiO_2 was 299 (kJ/mol). M. Guedes et al [13]

found that for samples containing a combination of TiO_2 and ZrO_2 nucleating agents, the Avrami constant, n , would change from 1 to 3 by varying the crystallization activation energy from 132 to 195.8 (kJ/mol). However Min and Hu et al [14-15] reported crystallization activation energy was varied from 303 to 425 (kJ/mol) and Avrami constant was 2.8 in a system with the same nuclei. Zheng et al [16] showed the activation energy for crystallization of a sample containing TiO_2 (2.3 wt %), ZrO_2 (2 wt %) and Y_2O_3 (4.46 wt%) was 512 (kJ/mol).

There are several methods for determining the crystallization mechanism in glass-ceramics such as Marotta [17], Matusita [18], Modified Kissinger [19] and Augis-Bennett [15].

In this research work, the crystallization mechanism and activation energy in $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glass ceramic systems, containing a mixture of TiO_2 , ZrO_2 , P_2O_5 , CeO_2 and Y_2O_3 (in a single, double, triple and quadruple nuclei forms) were investigated by using various methods through differential thermal analysis(DTA), Marotta and Augis-Bennett methods.

II. EXPERIMENTAL PROCEDURE

In this article a glass which is mentioned as S with the composition showed in table 1 was used as the base glass. In all other samples the letters T, Z, Y and C, show presence of TiO_2 , ZrO_2 , Y_2O_3 and CeO_2 respectively with the indication of their weight percent in front of them. The amount of added nuclei is also shown in table 1.

Sample	Oxides										
	SiO_2	Al_2O_3	Li_2O	Na_2O	K_2O	MgO	ZnO	TiO_2	ZrO_2	Y_2O_3	CeO_2
S	66.4	23.04	5.2	0.6	0.6	2.08	2.08	-	-	-	-
ST_3	66.4	23.04	5.2	0.6	0.6	2.08	2.08	3	-	-	-
ST_3Z_1	66.4	23.04	5.2	0.6	0.6	2.08	2.08	3	1	-	-
$\text{ST}_3\text{Z}_1\text{Y}_1$	66.4	23.04	5.2	0.6	0.6	2.08	2.08	3	1	1	-
$\text{ST}_3\text{Z}_1\text{C}_3$	66.4	23.04	5.2	0.6	0.6	2.08	2.08	3	1	-	3

Table1. Chemical composition of glasses (weight percentage)

The raw materials were reagent grade α - Al_2O_3 (PB-502 Alumina, Martinswerk Company, $d_{50}<45 \mu\text{m}$), SiO_2 (with purity up to 99%, $d_{50}<45 \mu\text{m}$) and Li_2CO_3 , NaCO_3 , K_2CO_3 , $\text{Mg}(\text{OH})_2$, ZnO , TiO_2 , ZrO_2 , P_2O_5 , Y_2O_3 and CeO_2 were purchased from Merck Company and used without further purification. The mixture of raw materials after mixing thoroughly were transferred to an alumina crucible and melted at 1650°C for 2 hours in an electric furnace. Afterwards the melt was cast in pre-heated stainless steel moulds and cooled naturally to room temperature. The thermal behavior of glass samples was monitored by DTA technique utilizing a simultaneous thermal analyzer (DTG-60 AH Shimadzu). Nucleation temperature (T_n),

crystallization peak temperature (T_p), crystallization activation energy (E), Avrami constant (n) and kinetic constant (m) were determined from DTA results. The reference material in these experiments was α - Al_2O_3 powder and heating rates were 10.5, 12.5, 15 and 17.5°C/min. Dilatometric softening temperatures (T_d) and Thermal Expansion Coefficient (TEC) were also measured by a dilatometer (model E-402 Netzsch). The optimum nucleation temperatures of samples were determined by Ray & Day method [20]. An X-ray diffractometer (Siemons-D500) was used in order to identify the produced crystalline phases in heat-treated samples. Cu- α radiation was used with a 20 kV emitter. The samples were polished and etched in 5%HF solution for 30 seconds and then coated with a thin film of gold for evaluation of particle shapes in samples by using a scanning electron microscope (LEO 440i). The lattice parameters of hexagonal β -Eucryptite solid solution were determined by using the following formula [8]:

$$d_{hkl}^2 = \frac{a^2}{\frac{4}{3}(h^2 + hk + k^2) + l^2 \frac{a^2}{c^2}} \quad (1)$$

In which d , (hkl) and (a,c) refer to distance between lattice planes (d -spacing), Miller index and lattice parameters.

III. RESULTS AND DISCUSSION

The base glass sample used in this article (S) is composed of SiO_2 , Al_2O_3 , Li_2O , Na_2O , K_2O , MgO and ZnO . The effect of these compounds on properties of silicate glass has been investigated by several researchers thoroughly. X. Guo, et al [11] has studied the effect of Na_2O , K_2O and MgO on silicate glass. They found that presence of Na_2O and K_2O in the base composition of the glass increases non-bridging oxygen ions and also accelerates the bulk nucleation and subsequent crystallization. However it was found that it decreases the viscosity. Similarly presence of MgO at high temperatures showed the same effect on viscosity. H. Bach, et al [21] has found that ZnO can improve the workability of silicate glass. In addition, it was found that replacement of ZnO with Li_2O decreases the thermal expansion coefficient. However replacement with MgO showed the opposite effect. Also he has reported the amount of Al_2O_3 must be in the range of 18-25wt% in order to provide the low thermal expansion coefficient, produce a transparent glass and also prevent Mullite phase formation. Furthermore he has found that the high amount of SiO_2 might increase the viscosity and cause non-homogeneity in this kind of glass ceramic.

In this article, in the first step, in order to investigate the effect of nuclei type and amount, several nuclei were added to the composition of the base glass in various weight percent. According to Ray and Day method [20] the sample with sharpest crystallization peak and highest intensity is mentioned as the optimized one in each group. This sharp and intense peak shows existence of high concentration of an effective nucleating agent and subsequently high value of growth rate. Also according to this method, for the two curves with the same sharpness, the one with lower peak temperature would be considered as the optimized one. The obtained results indicate that the sample containing 3wt % TiO_2 which is mentioned as ST_3 was the optimized sample. Fig 1 represents DTA curves of samples containing various

amount of TiO_2 . Each curve showed only one exothermic peak which is associated with precipitation of β -Eucryptite which is called high quartz solid solution (h-quartz).

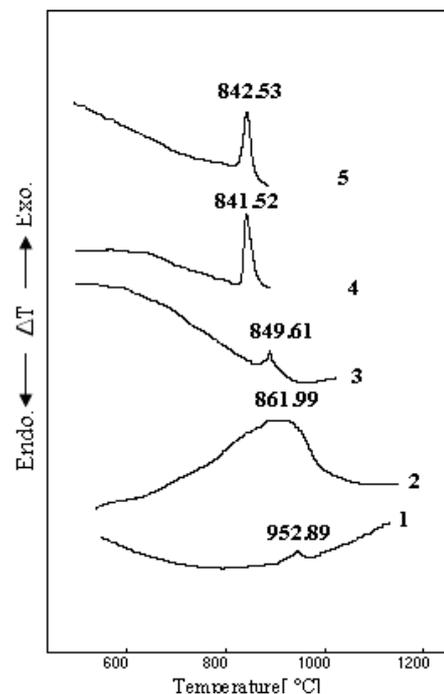


Fig.1 DTA curves of the glasses S(1), ST_1 (2), ST_2 (3), ST_3 (4) and ST_4 (5) at the heating rate of 10 (°C/min)

Presence of low concentrations of TiO_2 as a nucleating agent decreases the crystallization activation energy and thus causes the β -Eucryptite to β -spodumene transformation. At the same time it decreases the viscosity and melting point of the glass [22]. The obtained results showed that the crystallization temperature decreases from 925 to 841°C with increasing TiO_2 amount from 1 to 3 wt%, whereas increasing up to 4 wt% does not have a significant effect on this temperature. However it decreases the intensity of the crystallization peak. Also presence of TiO_2 as a nucleating agent causes phase separation. Doherty [23] showed that the phase separation occurred during cooling from the melt and subsequent heating, causes the formation of a large number of Aluminium Titanate crystals ($\text{Al}_2\text{Ti}_2\text{O}_7$) in SiO_2 -poor regions. These crystals act as sites for heterogeneous nucleation and h-quartz phase formation [21, 24]. So the peaks on DTA curves refer to the formation of h-quartz, which is confirmed by X-ray diffraction results. The same method was used for samples containing various amounts of ZrO_2 . DTA curves of these samples are represented in Fig 2. Similar to TiO_2 case, each curve shows an exothermic peak which its intensity reduces gradually with increasing the amount of ZrO_2 . The presence of ZrO_2 as a nucleating agent increases the viscosity of the melt as well as the crystallization activation energy, which shifts the peak temperature up to high values and also encourages the formation of β -Spodumene. So the sample containing 1wt % ZrO_2 which is mentioned as SZ_1 was chosen as the optimized sample in this group. Results showed that TiO_2 is more appropriate than ZrO_2 as a nucleating agent, because of the high power of ionic field of Ti^{4+} (1.25) in comparison with Zr^{4+} (0.78) and more ion absorption from

nearest neighboring atoms which increases the phase separation. Another problem with ZrO_2 oxides is their weak solubility in silicate melts which limits the amount of ZrO_2 that can be used as a nucleating agent to 3-4 Wt%. Furthermore, in comparison with TiO_2 , the melt contains ZrO_2 as a nucleating agent has a high value of viscosity. So the high viscosity of ZrO_2 in glass melt increases the crystallization activation energy and decreases the crystallization rate and subsequently shifts the peak temperature up to high values [25].

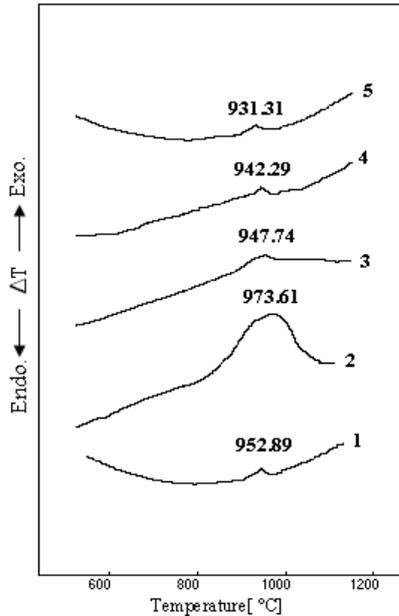


Fig.2 DTA curves of the glasses S (1), SZ_1 (2), SZ_2 (3), SZ_3 (4) and SZ_4 (5) at the heating rate of $10^\circ\text{C}/\text{min}$

The same method was used for samples containing P_2O_5 , Y_2O_3 and CeO_2 as single nucleating agents.

Results showed wide and indeterminate peaks for samples containing various amounts of P_2O_5 . Using this nucleating agent results in direct formation of β -Spodumene without transformation to β -Eucryptite and also decreasing crystallization temperature. For samples containing Y_2O_3 as a nucleating agent, just a sample with 1wt% Y_2O_3 showed the peak with low intensity which vanished with increasing the amount of this agent. Also the results did not show any sharp peak for samples containing CeO_2 as a nucleating agent.

Therefore our investigations showed that ST_3 (the sample contains 3Wt % TiO_2) was the best composition among the samples containing single nucleating agent.

In the second step ZrO_2 , P_2O_5 , Y_2O_3 and CeO_2 (1-4 wt %) were mixed with ST_3 and used as double (two-component) nucleating agents. According to DTA curves which are shown in Fig 3 and Ray and Day method, sample contained TiO_2 along with 1wt % ZrO_2 as nucleating agents (mentioned as ST_3Z_1), showed better crystallization activation energy.

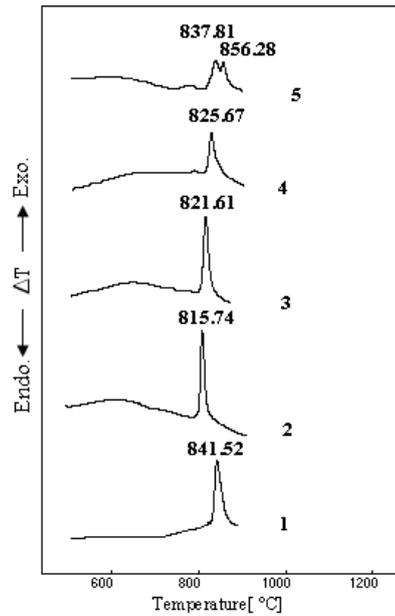


Fig.3 DTA curves of the glasses ST_3 (1), ST_3Z_1 (2), ST_3Z_2 (3), ST_3Z_3 (4) and ST_3Z_4 (5) at the heating rate of $10^\circ\text{C}/\text{min}$

It can be seen that the crystallization temperature decreases from 841°C for ST_3 to 815°C for ST_3Z_1 . However by increasing ZrO_2 amount, the temperature increases slowly up to just below the ST_3 peak temperature and the intensity of the peaks decreases. Two peaks can be seen for ST_3Z_4 . It is assumed that the first one refers to the formation of h-quartz and the second one refers to the transformation of h-quartz to β -Spodumene [26-27]. Hsu et al. [24] found that TiO_2 and ZrO_2 act as nucleating agents and cause phase separation by precipitating as $ZrTiO_4$ phase in SiO_2 -poor regions. As contrast, Maier et al. [3] studied a lithium aluminosilicate glass containing TiO_2+ZrO_2 . They also observed the formation of $ZrTiO_4$ crystallites which acted as precursor nuclei for subsequent crystallization. Fig 4 represents the phase separation in ST_3Z_1 glass ceramic.

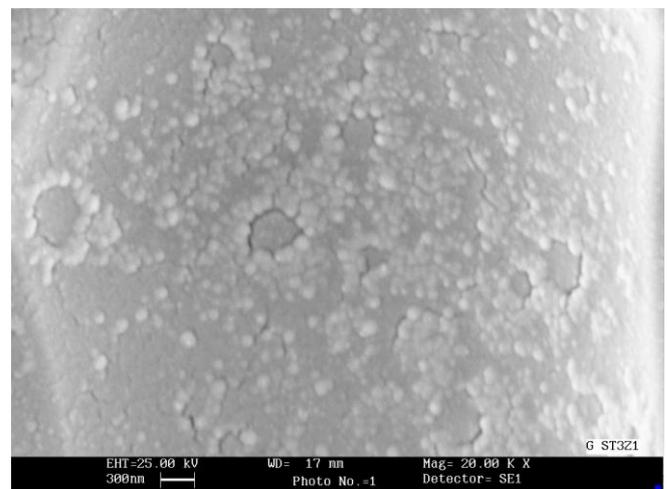


Fig.4 SEM micrograph of ST_3Z_1 glass that represents the separated regions in glass matrix

Results showed that the simultaneous use of TiO_2 and P_2O_5 as nucleating agents could not improve the nucleation rate and no sharp peak in DTA curves was observed. Also the

crystallization temperature increased and the intensity of the peak decreased by simultaneous use of TiO_2 and Y_2O_3 as double nucleating agents.

Hu et al [28] found that CeO_2 has a significant role in decreasing viscosity and encouraging the crystallization process as a flux in LAS glass-ceramics. They found that addition of CeO_2 not only lowers the viscosity of the glass, but also promotes crystallization. However it accelerates the transformations of glass to h-quartz and h-quartz to β -Spodumene. Our results showed that the presence of CeO_2 decreased the intensity of the crystallization peak so TiO_2 - CeO_2 couple is not appropriate as a double nucleating agent in LAS glass ceramics.

However our investigations showed that ST_3Z_1 (the sample contains 3Wt% TiO_2 and 1Wt% ZrO_2) was the best composition between samples contain double nucleating agent.

In the third step P_2O_5 , Y_2O_3 and CeO_2 (1-4 wt %) were added to ST_3Z_1 in order to investigate the effect of triple nucleating agents on crystallization mechanism of LAS glass ceramics. P_2O_5 didn't show an appropriate result in these series. It seems to be because of its electrical neutralizing effect on Al^{3+} in AlO_4 tetrahedral. According to the results, addition of 1wt% Y_2O_3 ($\text{ST}_3\text{Z}_1\text{Y}_1$) and 3wt% CeO_2 ($\text{ST}_3\text{Z}_1\text{C}_3$) to ST_3Z_1 provides better bulk nucleation and crystallization, than other samples in this group.

In the fourth step, P_2O_5 and CeO_2 were added to $\text{ST}_3\text{Z}_1\text{Y}_1$ and $\text{ST}_3\text{Z}_1\text{C}_3$ (1-4wt %) in order to investigate the effect of quadruple nucleating agents on crystallization mechanism of these glass ceramics. In both cases crystallization temperature increased (above 900°C). Also the sharpness of exo-peaks decreased gradually and eventually the peaks vanished. Therefore the results showed the samples contained quadruple nucleating agents were not appropriate.

ST_3Z_1 , $\text{ST}_3\text{Z}_1\text{Y}_1$ and $\text{ST}_3\text{Z}_1\text{C}_3$ were chosen as optimized samples. They were kept for 3 hours at several temperatures above their glass transition temperature (T_g) i.e. 600 - 735°C for ST_3Z_1 , 684 - 754°C for $\text{ST}_3\text{Z}_1\text{Y}_1$ and 690 - 740°C for $\text{ST}_3\text{Z}_1\text{C}_3$. Then the nucleation temperature (T_n) and crystallization temperature (T_p) was determined by Ray and Day method according to DTA results of each sample which are shown in figures 5 to 7.

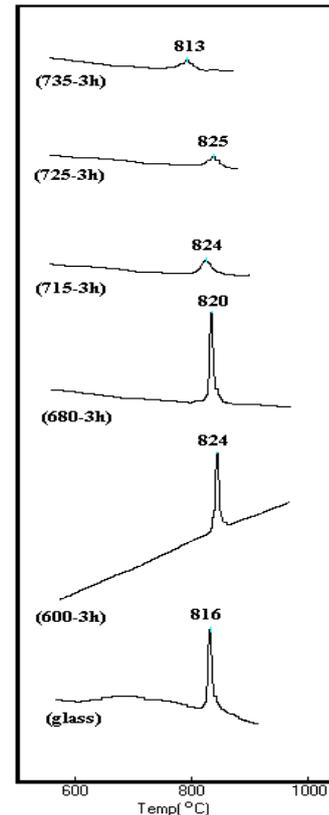


Fig.5 DTA plots of ST_3Z_1 glass heat treated at different temperatures with soaking time for 3 hours (in order to determine the nucleation temperature)

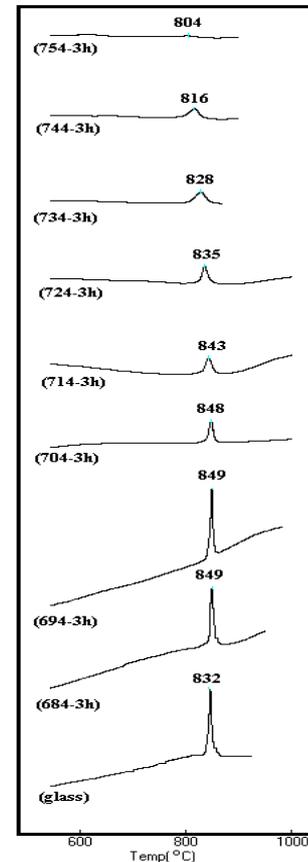


Fig.6 DTA plots of the $\text{ST}_3\text{Z}_1\text{Y}_1$ glass heat treated at different temperatures with soaking time for 3 hours (in order to determine the nucleation temperature)

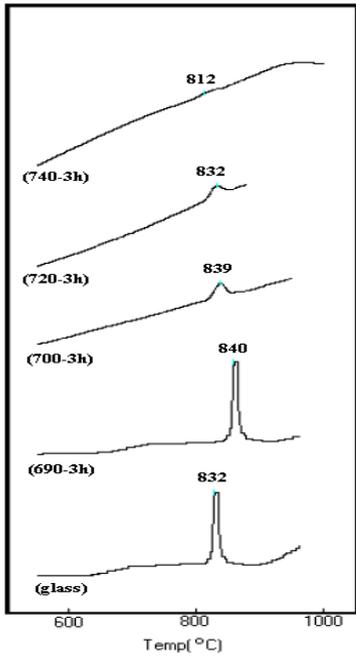


Fig.7 DTA plots of the $ST_3Z_1C_3$ glass heat treated at different temperatures with soaking time for 3 hours (in order to determine the nucleation temperature)

According to the results the optimum nucleation temperatures are 680, 694 and 690°C and the optimum crystallization temperatures at these nucleation temperatures are 820, 849 and 840°C for ST_3Z_1 , $ST_3Z_1Y_1$, $ST_3Z_1C_3$, respectively.

Fig 8 represents the XRD patterns of the three optimized samples after 3 hours of heat treatment at their crystallization temperature.

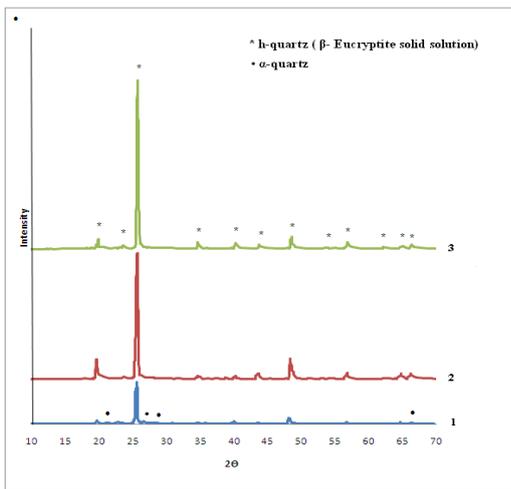


Fig.8 X-ray diffraction pattern for $ST_3Z_1C_3$ (1), $ST_3Z_1Y_1$ (2) and ST_3Z_1 (3) at their DTA peak crystallization temperature for 3 h

As it can be seen, main crystalline phase in all samples is h-quartz (β - Eucryptite). However free quartz (α -quartz) was detected beside the main phase for $ST_3Z_1C_3$. It has been known that formation of free quartz leads to higher Thermal Expansion Coefficients (TEC) and lower

thermal shock resistances in this kind of glass ceramics [29].

As it is presented in XRD results, the intensities of h-quartz peaks decrease by adding Y_2O_3 and CeO_2 to ST_3Z_1 .

The lattice parameters of h-quartz (β -eucryptite solid solution) have been calculated from the measured d-value of crystal plane. The crystal planes of (101), (112) and (100) have been used for determination of lattice parameters which are shown in Table 2. These results are comparable with lattice constants of h-quartz-type and Keatite-type aluminosilicates ($Zn_{0.5}AlSi_2O_6$ and $LiAlSi_2O_6$ composition) [21]. Therefore, it can be deduced that the optimized samples should have a low thermal expansion coefficient of about $1.7 \times 10^{-6} / ^\circ C$ because of their approximately equal lattice constants to the above-mentioned compositions. Thermal expansion coefficients of the optimized samples were measured by dilatometer test between room temperature (T_R) and dilatometric softening temperature (T_d).

Table2. Unit-cell dimension for optimum samples and various h-quartz type alumina-silicates

Composition	Lattice Constants, °A	
	a	c
ST_3Z_1	5.200	5.434
$ST_3Z_1Y_1$	5.221	5.661
$ST_3Z_1C_3$	5.208	5.850
$LiAlSi_2O_6$	5.212	5.457
$Zn_{0.5}AlSi_2O_6$	5.220	5.460

Coefficient of linear thermal expansion – the linear thermal expansion per temperature change- is represented in the following equation (2):

$$\alpha = \frac{\Delta L / L_1}{\Delta T} = \frac{L_2 - L_1}{L_1(T_2 - T_1)} \quad (2)$$

Where L_1 and L_2 are the lengths of the specimen at the test temperatures of T_1 and T_2 [30]. The dilatometric curves of these samples are shown in Fig 9. Dilatometric softening temperature (T_d), glass transition temperature (T_g) and thermal expansion coefficient (α) for the optimized samples are summarized in table 3.

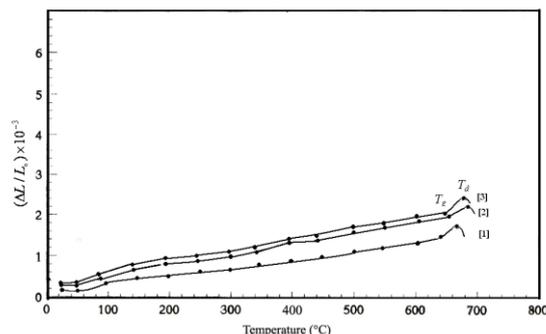


Fig.9 Dilatometric curves for ST_3Z_1 (1), $ST_3Z_1Y_1$ (2) and $ST_3Z_1C_3$ (3)

Table3. Dilatometric softening temperature (T_d), glass transition temperature (T_g) and thermal expansion coefficient (α) for the optimized samples

Sample Name	$T_d(^{\circ}\text{C})$	$T_g(^{\circ}\text{C})$	$\alpha (\times 10^{-6}/^{\circ}\text{C})$
ST ₃ Z ₁	640	671	1.65
ST ₃ Z ₁ Y ₁	656	686	1.73
ST ₃ Z ₁ C ₃	650	680	1.93

Thermal expansion coefficient of ST₃Z₁ showed the smallest amount among these samples which is about $2.2 \times 10^{-6} / ^{\circ}\text{C}$ smaller than this amount for LAS glass ceramics.

Different methods such as Marotta and Augis-Bennett were used in order to investigate the crystallization mechanism of the optimized samples.

Fig 10 and 11 show the results of variation of $\text{Ln} \Delta T$ versus $1/T$ and $\text{Ln} \alpha$ versus $1/T_p$ for ST₃Z₁Y₁ sample derived from Marotta equations (equation 3 and 4) [17, 31, 28]:

$$\text{Ln}T = \frac{-nE}{RT} + \text{const.} \quad (3)$$

$$\text{Ln} \alpha = \frac{-E}{RT_p} + \text{const.} \quad (4)$$

Where α , T_p , E , R , n and ΔT indicate the heating rate, crystallization temperature, crystallization activation energy, gas constant, the value of Avrami constant and temperature deviation from the baseline, respectively. The plots of both patterns are expected to be linear, and value of E and n can be measured from the slope of the two lines.

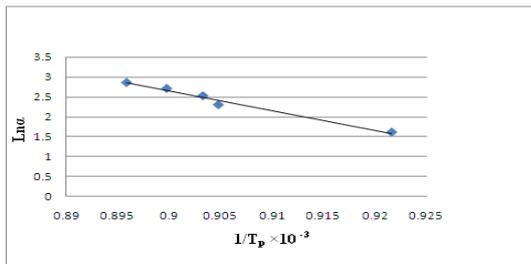


Fig.10 Variation of $\text{Ln} \alpha$ vs $1/T_p$ in ST₃Z₁Y₁ for determination of the crystallization activation energy according to Marotta method

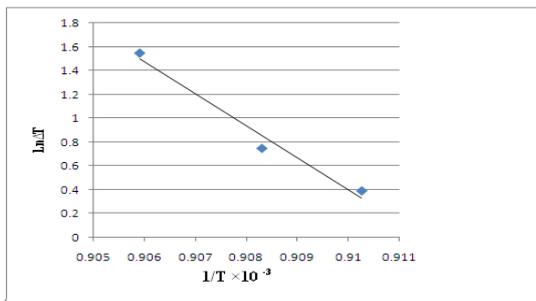


Fig.11 Variation of $\text{Ln} \Delta T$ vs $1/T$ in ST₃Z₁Y₁ for determination of the Avrami exponent according to Marotta methods

The crystallization kinetic parameters of the optimized samples can also be calculated by Augis-Bennett equations (equation 5 and 6) which are expressed as below [15]:

$$\text{Ln} \frac{T_p^2}{\alpha} = \frac{E}{RT_p} + \text{const.} \quad (5)$$

$$n = \frac{2.5}{\Delta T} \times \frac{RT_p^2}{E} \quad (6)$$

In this equation, ΔT is the width of the exothermic peak at the half maximum intensity. The value of n close to 1 implies a surface crystallization mechanism, n close to 2 refers to one-dimensional crystallization, n close to 3 implies a two-dimensional bulk crystallization process and the value of n close to 4 implies three-dimensional bulk crystallization [32].

The plot of $\text{Ln} \frac{T_p^2}{\alpha}$ vs. $1/T_p$ is shown in Fig 12.

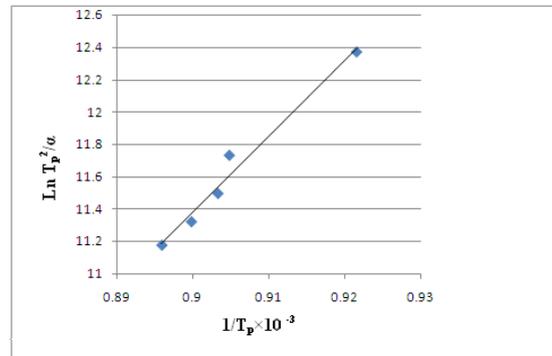


Fig.12 The plots of $\text{Ln} \frac{T_p^2}{\alpha}$ vs. $1/T_p$ for ST₃Z₁Y₁ for

determination of the crystallization activation energy according to Augis-Bennett method.

It is used for calculation of the crystallization activation energy and Avrami constant, n , of the optimized samples which are summarized in Table 4.

Table4. Avrami exponent and crystallization activation energy determined by various methods

Sample	Avrami Constant (n)		Activation Energy (E) (kJ/mol)	
	Marotta	Augis-Bennett	Marotta	Augis-Bennett
ST ₃ Z ₁	4.23	4.38	269.11	255.2
ST ₃ Z ₁ Y ₁	3.70	3.81	287.8	273.39
ST ₃ Z ₁ C ₃	2.96	3.03	365.6	349.67

The n values determined from the Marotta and Augis-Bennett methods were almost identical. Considering the amount of the Avrami constants were calculated for the optimized samples, it can be deduced that in ST₃Z₁ and ST₃Z₁Y₁ with n value of about 4, the crystallization

mechanism were three-dimensional which indicates homogeneous crystallization. Therefore the shapes of the crystalline particles in these samples are spherical. Comparison of E values (crystallization activation energy) in ST_3Z_1 and $ST_3Z_1Y_1$ indicates the same predominate crystallization mechanism. However, it can be seen that Y_2O_3 increases the crystallization activation energy, which would probably leads to a reduction in crystallization rate in $ST_3Z_1Y_1$ glass ceramics. Comparison of n and E values in ST_3Z_1 and $ST_3Z_1C_3$ glass ceramics indicates that addition of CeO_2 to the composition increases the crystallization activation energy. Therefore, considering the amounts of the Avrami constants calculated from the Marotta and Augis-Bennett methods in $ST_3Z_1C_3$ sample, the crystallization mechanism was two-dimensional bulk type and the shapes of the particles were plate like. It should be noted that in spite of the high value of n, the E value was very low, which seems to be the lowest reported for this system [28].

Fig.13 shows the SEM micrograph of $ST_3Z_1Y_1$ and $ST_3Z_1C_3$ glass ceramics which have been nucleated at T_n (optimized nucleation temperature) and crystallized at T_p (optimized crystallization temperature) for 3 hours.

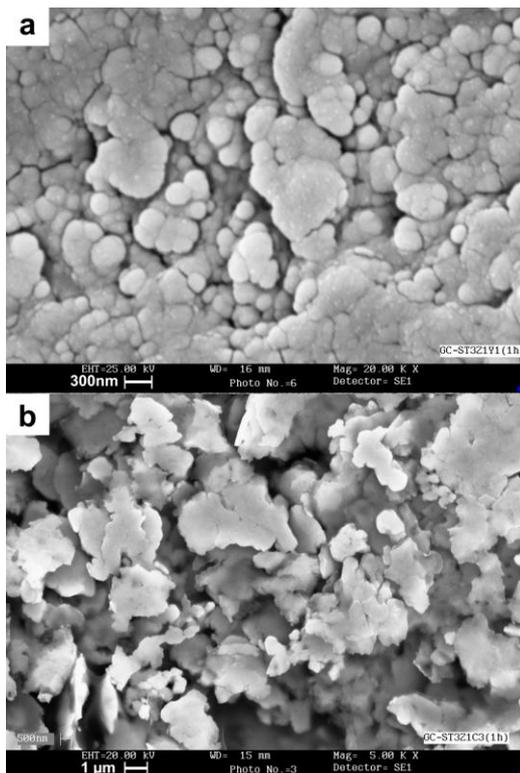


Fig.13 SEM micrograph of of $ST_3Z_1Y_1$ (a) and $ST_3Z_1C_3$ (b) nucleated at T_n and crystallized at T_p for 3 h

The presence of plate-like and spherical crystalline particles in morphologies of $ST_3Z_1C_3$ and $ST_3Z_1Y_1$ can be seen respectively, which is again an evidence for a two dimensional and three dimensional crystallization mechanisms in these samples respectively. As it can be seen, the precipitated crystalline particles of $ST_3Z_1Y_1$ are smaller than 300 nm and $ST_3Z_1C_3$ particles are bigger than 1 μm . The fine texture of $ST_3Z_1Y_1$ which was produced from a suitable nucleating agent would lead to a high bending

strength in this sample. According to the above-mentioned discussion these glass-ceramics can be used as high thermal shock resistance products for commercial applications.

IV. CONCLUSIONS

According to DTA results, simultaneous use of TiO_2 , ZrO_2 and CeO_2 as nucleating agents with various ratios is a proper approach to obtain a high amount of the crystalline phase in bulk crystallization of glass-ceramics in $Li_2O-Al_2O_3-SiO_2$ systems. The XRD results revealed that the samples are composed of free quartz as the minor phase and h-quartz (β -Eucryptite) as the main phase. The most suitable nucleation temperatures of the optimized samples were 680, 694 and 690°C for ST_3Z_1 , $ST_3Z_1Y_1$ and $ST_3Z_1C_3$, respectively. The kinetic parameters (m,n) derived from Marotta and Augis-Bennett equations, showed that glass-ceramics containing both ZrO_2 and TiO_2 nuclei, represent bulk crystallization. Also samples containing ZrO_2 , TiO_2 and CeO_2 in the triple nuclei forms, represent two-dimensional bulk crystallization, which were confirmed by SEM analysis of microstructures of these samples. The minimum crystallization activation energy and maximum Avrami constant were obtained by combination of 3Wt% TiO_2 and 1Wt% ZrO_2 as nucleating agents.

V. ACKNOWLEDGMENT

The authors gratefully acknowledge Mr. Nad Ali Arefian for his helpful feedback on the manuscript.

VI. REFERENCES

1. McMillan P W. Glass-ceramics. 2nd ed. Non-metallic solids. London ; New York: Academic Press; 1979.
2. Strnad Z E. Glass-ceramic materials : liquid phase separation, nucleation, and crystallization in glasses. Glass science and technology v. 8. Amsterdam ; New York: Elsevier; 1986.
3. Mueller G, Maier V. Mechanism of oxide nucleation in lithium aluminosilicate glass-ceramics. J Am Ceram Soc. 1987;70:C176-C8.
4. Hsu J Y, Speyer R F. Influences of zirconia and silicon nucleating agents on the devitrification of $Li_2O-Al_2O_3-6SiO_2$ glass. J Am Ceram Soc 1990;73:3583-93.
5. James P F. Glass ceramics: new compositions and uses. Journal of Non-Crystalline Solids. 1995;181(1-2):1-15.
6. Barbieri L, Corradi A B, Leonelli C, Manfredini T, Romagnoli e M, Siligardi C, Mustarelli P, Tomasi C. Nucleation and Crystallization of a Lithium Alumino Silicate Glass. J Am Ceram Soc 1997;80(12):3077-83
7. Beall G H, Beall L R G H , Pinckney L R. Nanophase Glass-Ceramics J Am Ceram Soc 1999;85:5-16
8. Arnault L, Gerland M. Microstructural Study of two LAS-type Glass-Ceramics and Their Parent Glass. J Mat Sci. 2000;35:2331-45.
9. Riello P, Canton P, Comelato N, Polizzi S, Verità M, Fagherazzi G et al. Nucleation and crystallization behavior of glass-ceramic materials in the $Li_2O-Al_2O_3-SiO_2$ system of interest for their transparency properties. Journal of Non-Crystalline Solids. 2001;288(1-3):127-39.

10. Armakar B, Kundu P, Jana S. Crystallization kinetics and mechanisms of low-expansion lithium-alumino silicate glass-ceramics by dilatometry. *J Am Ceram Soc.* 2002;85(10):2572-4.
11. Guo X, Yang H, Cao M. Nucleation and crystallization behavior of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ system glass-ceramic containing little fluorine and no-fluorine. *Journal of Non-Crystalline Solids.* 2005;351(24-26):2133-7.
12. Sung Y-M, Dunn SA, Koutsky JA. The effect of boria and titania addition on the crystallization and sintering behavior of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-4\text{SiO}_2$ glass. *Journal of the European Ceramic Society.* 1994;14(5):455-62.
13. Guedes M, Ferro AC, Ferreira JMF. Nucleation and crystal growth in commercial LAS compositions. *Journal of the European Ceramic Society.* 2001;21(9):1187-94.
14. An-Min H, Kai-Ming L, Fei P, Guo-Liang W, Hua S. Crystallization and microstructure changes in fluorine-containing $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glasses. *Thermochimica Acta.* 2004;413(1-2):53-5.
15. Hu AM, Liang KM, Li M, Mao DL. Effect of nucleation temperatures and time on crystallization behavior and properties of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glasses. *Mater Chem Phys.* 2006;98(2-3):430-3. doi:10.1016/j.matchemphys.2005.09.060.
16. Zheng W, Cheng J, Tang L, Quan J, Cao X. Effect of Y_2O_3 addition on viscosity and crystallization of the lithium aluminosilicate glasses. *Thermochimica Acta.* 2007;456(1):69-74.
17. Marotta A, Saiello S, Branda F, Buri A. Activation energy for the crystallization of glass from DDTA curves. *J Mater Sci.* 1982;17:105-8.
18. Matusita K, Sakka S. Study on crystallization kinetics in glass by differential thermal analysis. *Thermochimica Acta.* 1979;33(0):351-4.
19. Matusita K, Sakka S. Kinetic study of crystallization of glass by differential thermal analysis—criterion on application of Kissinger plot. *Journal of Non-Crystalline Solids.* 38-39, Part 2(0):741-6.
20. Ray C S, Fang X, Day D E. New Method for Determining the Nucleation and Crystal-Growth Rates in Glasses. *J Am Ceram, Soc.* 2000;83(4):865-72.
21. Bach H. Low thermal expansion glass ceramics. Schott series on glass and glass ceramics. Berlin ; New York: Springer; 1995.
22. Al-Harbi O A, Esmat M A, Hamzawy, Mujtaba Khan M. Effect of Different Concentrations of Titanium Oxide (TiO_2) on the Crystallization Behavior of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ Glasses Prepared from Local Raw Materials. *Journal of Applied Sciences.* 2009;9(16):2981-6.
23. Doherty P E, Lee D W, Davis R S. Davis. Direct Observation of the Crystallization of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ Glasses Containing TiO_2 . *J Am Ceram Soc.* 1967;52(2):77-81.
24. Hsu A M, Liang K M, Wang G, Zhou F, Peng F. Effect of nucleating agents on the crystallization of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ system glass. vol 3. Dordrecht, PAYS-BAS: Springer; 2004.
25. Nocun M, Baugajski W. Effect of Y-PSZ Additive on the Structure, Thermal Behavior and Mechanical Properties of β -Spodumene Ceramics. Aedermannsdorf, SUISSE: Trans Tech; 1997.
26. Barry TI, Clinton D, Lay LA, Mercer RA, Miller RP. The crystallisation of glasses based on the eutectic compositions in the system $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$. *Journal of Materials Science.* 1970;5(2):117-26.
27. Coon DN, Neilson RM. Effect of MgO Addition on the Glass Transition Temperature of LAS Glasses. *Journal of Materials Science Letters.* 1988;7(1):33-5.
28. Hu AM, Liang KM, Zhou F, Wang GL, Peng F. Phase transformations of $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glasses with CeO_2 addition. *Ceramics International.* 2005;31(1):11-4.
29. García-Moreno O, Fernández A, Torrecillas R. Solid state sintering of very low and negative thermal expansion ceramics by Spark Plasma Sintering. *Ceramics International.* 2011;37(3):1079-83.
30. Serbena FC, Soares VO, Peitl O, Pinto H, Muccillo R, Zanutto ED. Internal Residual Stresses in Sintered and Commercial Low Expansion $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ Glass-Ceramics. *Journal of the American Ceramic Society.* 2011;94(4):1206-14. doi:10.1111/j.1551-2916.2010.04220.x.
31. Knickerbocker S, Tuzzolo MR, Lawhorne S. Sinterable β -Spodumene Glass-Ceramics. *Journal of the American Ceramic Society.* 1989;72(10):1873-9.
32. Matusita K, Komatsu T, Yokota R. Kinetics of non-isothermal crystallization process and activation energy for crystal growth in amorphous materials. *Journal of Materials Science.* 1984;19(1):291-6.

Mining Two Class Opinions Using Optimized Recurrent Neural Network

J. Isabella, R. M. Suresh,

Research Scholar, Sathyabama University, Chennai, India
Principal, Jerusalem College of Engineering, Chennai, India

Abstract: Opinion mining automatically classifies the sentiment of the reviews of the customers into positive or negative. Sentiment classification helps analyzing the customer's opinions from the information gathered online. Opinions of products and services are available at various sources such as feedback at websites, blogs in the internet. Automatic opinion mining is required as it is not possible to manually classify the amount of review available online. In this paper, it is proposed to classify movie reviews using a novel Recurrent Neural Network with genetic algorithm optimization.

Keywords: Opinion Mining, Sentiment Classification, Neural Network, Genetic algorithm

I. Introduction

With the emergence of Internet, number of people participating in posting opinions, feedbacks has drastically increased. These reviews are very useful for users to take decisions. Consumers tend to decide whether to buy a product or not based on the opinion. Reviews of products, movie, books, events, and political issues and so on, are available in abundance but it is not possible to go through all the content. So, opinion mining is used for creation and automatically updated review websites, which summarizes opinions on the whole of a particular product or event. Sentiment classification classifies the reviews into positive or negative opinions [1]. An individual finds it useful to see summary of opinions of existing users before making an informed decision. User can also compare with that of competing products. Reviews of movies, books, music are other oft-looked up issues.

Opinion mining uses information retrieval techniques for the process of searching, retrieving relevant documents for the information required. For efficient information retrieval, documents are transformed into a standard format and indexed. Stopword list is prepared which lists the words that are not relevant for retrieving documents. Common words are reduced to its stem or root word. Term frequency is computed, which expresses the number of times a term occurs in a document. Relative term frequency is obtained by term frequency to the total number of occurrence of all terms in the document. Similar documents have similar relative frequencies. Thus, using a similarity measure like cosine measure, documents similar to the query is retrieved. With the advance in machine learning, opinions and reviews of various products and services are semi-automatically or automatically retrieved and further classified. Several works in literature deal with mining reviews of automobiles, banks, movies, travel destinations, electronics and mobile devices [2].

Movie review mining has some special challenges when compared with product review mining. As movie review mining is very domain specific and word semantics in review could contradict with overall sentiment polarity (good or bad) of that review [3]. For example, an "unpredictable" service gives negative meaning to product review, whereas for a movie with "unpredictable" storyline gives positive opinion to moviegoers. Machine learning methods and semantic orientation methods are the main approaches used for sentiment classification. Bo Pang et al., [4] investigated the effectiveness of classification of documents by overall sentiment using machine learning techniques. Experiments showed that the machine learning techniques give better result than human produced baseline for sentiment analysis on movie review data. In this paper, it is proposed to classify movie reviews using Recurrent neural network with genetic algorithm.

II. Previous Research

Goldberg, et al., [5] presented a graph-based semi-supervised learning algorithm for rating the movie reviews. In the proposed method, the problem of the scarcity of labeled movie reviews to infer rating is addressed. A graph-based semi-supervised learning is proposed for learning from a set of movie reviews with ratings. The unlabeled documents are rated on the basis of the sentiment expressed in the review. This was done by creating a graph on both labeled and unlabeled data to encode certain assumptions for rating the reviews. The optimization problem is then solved to obtain a smooth rating function over the whole graph. Experimental results show that the proposed method achieved better predictive accuracy when compared with other methods with limited labeled data.

Dave, et al., [6] presented a classifier for sentiment classification of product reviews. Information retrieval techniques were used for feature extraction and scoring of words. Product reviews are obtained from the sites or clipping services. Structured reviews are used for training and testing to determine sentiment. The classifier is then used to identify and classify review sentences. Noise and ambiguity limits the performance, so sentences are grouped into attributes for better results. Experiments show that the method works as well as or better than traditional machine learning.

Ye, et al., [7] proposed a semantic approach for sentiment classification of movie reviews in Chinese. Word segmentation was introduced for classification process and also an optimal reference-word-pair selecting process for determining extremely positive and negative opinions. A semantic orientation (SO) value is calculated for the phrases. The average SO value of the review is then calculated to classify it as positive if the value exceeds the threshold and negative otherwise. Results show that the proposed method is comparable with the figures of other movie classification studies.

Zhuang, et al., [8] proposed an approach for movie review mining and summarization by integrating WordNet, and statistical analysis. In the proposed approach, a key word list is generated based on WordNet, movie and labeled training data to find features and opinions. Polarity of the opinion word is determined. Suitable feature-opinion pairs are identified by applying grammatical rules. Summaries of the reviews are generated using the extracted feature-opinion pairs. IMDB dataset was used for experimentation and the results showed the proposed method performed better than the other review mining algorithms.

Inferences made by Pang et al., [4] are that machine learning techniques are better than human baselines for sentiment classification. Whereas the accuracy achieved in sentiment classification is much lower when compared to topic based categorization. The experimental setup consists of movie-review corpus with randomly selected 700 positive sentiment and 700 negative sentiment reviews. Features based on unigrams and bigrams are used for classification. Learning methods Naïve Bayes, maximum entropy classification and support vector machines were employed. Machine learning algorithms using unigrams got better accuracy than human selected unigram. Whereas in topic based classification using bag of unigram features achieve 90% accuracy suggesting that sentiment categorization is more difficult than topic categorization.

Peter Turney [9] proposes an unsupervised learning algorithm, using semantic orientation of the phrases containing adjectives and adverbs, to classify reviews. The approach initially extracts phrases containing adjectives and adverbs; the semantic orientation of the phrase is estimated using PMI-IR; based on the average semantic orientation the phrases the review is classified as recommended (Thumbs up) or not recommended (Thumbs down). Experiment was conducted using 410 reviews on various topics; an average accuracy of 74% was achieved. Movie reviews had accuracy of 66% whereas banks and automobiles attained 80% to 84% accuracy. The advantage of this algorithm is its simplicity of using PMI-IR. The time required for queries is the main disadvantage of this method.

III. Research Method

Bo Pang and Lillian Lee have created a collection of movie-review documents from the Internet Movie Database (IMDb) archives. It is an online database of information related to movies, television shows, actors. In this paper, it is proposed to use online movie reviews as data. The dataset is labeled with respect to their overall sentiment polarity (positive or negative) or subjective rating (e.g., two stars). List of stop words for commonly occurring words and stemming words with similar context is prepared. A total of two hundred reviews with 100 positive and 100 negative are chosen in this work.

3.1 Recurrent Neural Network (RNN)

Neural networks are an artificial representation of human brain, artificial neurons use mathematical or computational mode for information processing. Neurons exhibit complex global behavior due to the connections between the neurons and its parameters. Neural networks are adaptive in nature, changing according to the information flow through the network. Neural networks learn by examples, during learning the connections between the neurons are adjusted. The neural network is made up of layers; input layer, hidden layer and output layer. The connections between the neurons between layers are referred to as weights. The inputs are multiplied by the weights and then fed through a transfer activation function to generate an output. During training, the weights of each neuron are so adjusted to reduce the error between the desired output and actual output. Back propagation algorithm [10] is the most popular learning algorithm used to train the neural networks.

Recurrent neural networks (RNNs) are dynamical systems used for both classification and prediction using the temporal information of the inputs. The interrelations between the inputs and the internal state are exploited to produce output during training of RNNs. Thus, the outputs represent the relevant internal states of the past [11]. During training process, a second source of information helps constitute the target values which specify the relevant interrelations in the input sequence. Inputs for RNNs are generally in a time series and the target is a time series or a sequence of constant value. RNNs when used for classification especially in language learning, the output are given in form of a constant class label whereas in prediction tasks the output consists of time series.

The neurons activations in the RNN depend on the input series and also the internal dynamics of the RNN. Thus, on training of the weights, the output neurons can classify an input sequence. Generally, a time series prediction task is included in the training as it improves the classification ability. This also helps to extract the best features from the input series. Topology of an extended Elman-network [11] is illustrated in the figure 1. The network consists of two parts; a feed-forward part and a memory part. The memory part stores the activation of the neurons of the previous cycle and forwards it as additional inputs.

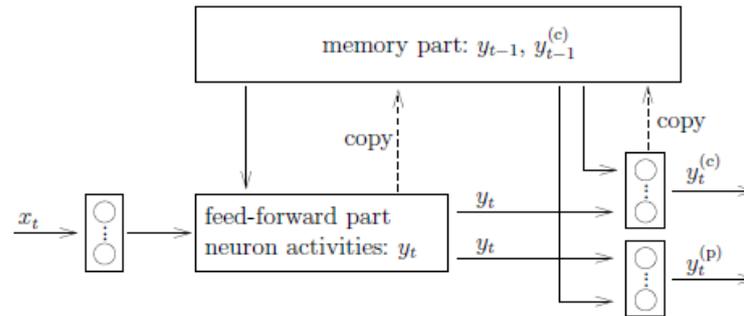


Figure 1: Topology of an extended Elman-network

The induced local field y_k of the neuron is given by equations as follows:

$$y_k = b_k + \sum_i \sum_j w_{kij} x_i u_j$$

$$x_k(n+1) = \varphi(v_k(n))$$

$$= \frac{1}{1 + \exp(-v_k(n))}$$

where b_k is the associated weight

$x_k(n)$ is the state of neuron k,

$u_j(n)$ is the input applied to source j

w_{kij} is the weight of neuron k

For learning in simple recurrent networks, error at state layer $\delta_j(t)$, is used to modify weights and errors are backpropagated according to equation 5.13

$$\delta_{ij}(t-1) = \sum_h^m \delta_{ih}(t) u_{hj} f'(y_{ij}(t-1))$$

Where h is the index for the activation receiving node and j for the sending node (one time step back). This allows calculation of the error as assessed at time t , for node outputs (at the state or input layer).

Though backpropagation is well suited for simple training problems, with increase in complexity the performance is reduced significantly. This is due to the fact that gradient search techniques get trapped at local minima. To overcome this problem, it is proposed to use genetic algorithm optimization for learning of neural network.

3.2 Genetic Algorithm

Genetic algorithms are nature-inspired optimization methods that could be advantageously used for optimization problems. GAs imitates basic principles of life and applies genetic operators like mutation, crossover, or selection to a sequence of alleles which is the equivalent of a chromosome in nature and is built by a representation which provides a string of symbols to every possible solution of the optimization problem. Earlier work revealed that behavior and performance of GEAs was strongly influenced by the representation used. Due to this many recommendations for a proper design of representations was made over the years. But most design rules are of a qualitative nature and are not very helpful for estimating how different representation types influence problem difficulty.

Genetic algorithms (GA) are algorithms used for optimization and learning. In GA, the solutions are encoded on chromosomes. Fitness function is used to evaluate the chromosomes. A population of chromosomes or solutions is initiated, and GA operators such as reproduction, mutation, and crossover are applied to generate the next generation of chromosomes. The population reproduces until stopping criteria is met or for a specified number of iterations. During each iteration, one or more parents are chosen to reproduce. The selection of parents depends on the fitness values, the chromosomes with high fitness value are chosen often to reproduce. Children are produced and inserted into the population. Thus, GA produces population of better and better solutions, converging towards global optimum.

The pseudo-algorithm of GA is as follows:

BEGIN

1. Generation t=0;
2. Initialize the population P(t);
3. While not Termination criterion do
 - a. Evaluate the population P(t);
 - b. Next P'(t) (by crossover/mutation) over P(t);
 - c. Evaluate P'(t);
 - d. P(t+1) = select from P'(t);
 - e. Generation t=t+1;
4. End while.

IV. Experimental Setup

It is proposed to investigate the effectiveness of GA to evolve the number of neurons in the hidden layer of the Neural Network architecture. The GA architecture used to optimize the number of hidden neurons is shown in Table 1.

Table 1: Parameters used in the GA optimization

Number of epoch	200
population size	10
Maximum generations	5
Neuron optimization lower bound	10
Neuron optimization upper bound	30
Encoder mechanism	Roulette
Cross over type	Two point
cross over probability	0.9
Mutation	Uniform
Mutation probability	0.01

The parameters used in the experiment are shown in Table 2. The table gives the details of the number of input neurons, number of neurons in the hidden layer and the momentum.

Table 2: The parameters used in proposed neural network design

Number of neurons in input layer	48
Number of neurons in output layer	2
Number of hidden layer	1
Context unit time constant	0.8 second
Transfer function of context unit	Integrator
Number of neurons in hidden layer	10, 15, 20, 25
Learning rule	Back propagation
Number of epochs for termination	500

V. Results

The classification accuracy obtained is compared with existing feed forward network and recurrent neural network. Figure 2 shows the classification accuracy of the three neural network techniques.

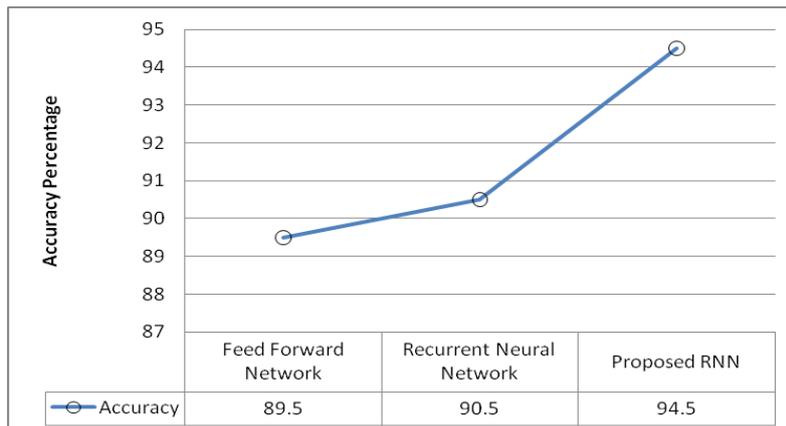


Figure 2: The classification accuracy compared with different classifiers

The plot of Avg MSE vs Epoch (number of iterations) for various number of hidden neurons is shown in Figure 3.

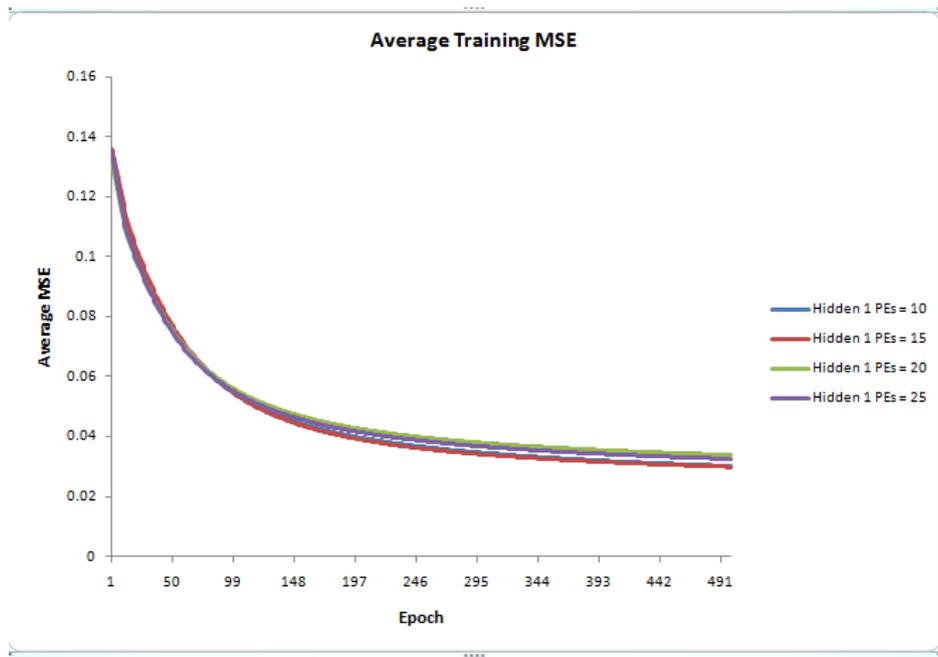


Figure 3: Avg MSE versus No.of Epoch

The best fitness obtained and average fitness obtained is shown in Table 3. Figure 3 shows the plot of MSE with the generation.

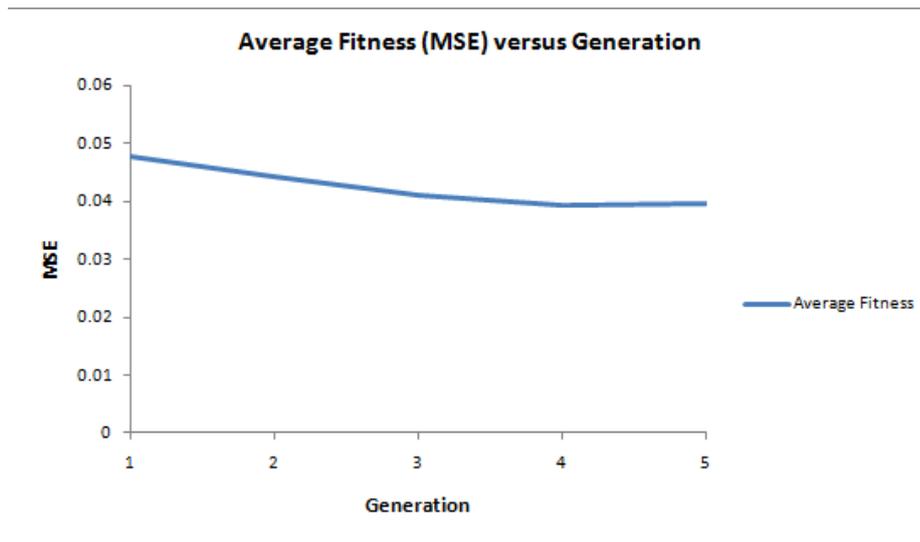


Figure 4: The average fitness value obtained

Table 3: Fitness values obtained

Best Fitness	Average Fitness
0.04174498	0.04770372
0.03654528	0.0443747
0.03654528	0.04110919
0.03654528	0.03932239
0.03654528	0.03977443

VI. Summary and Concluding Remarks

In this work an improved Recurrent Neural Network with Genetic Optimization was proposed. Features were extracted from the unstructured movie review data and feature reduction was done based on correlation between features. The proposed classifier algorithm shows an improvement of 4.42 % over conventional Recurrent Neural Network algorithm. Using genetic optimization, the ideal number of neurons in the hidden layer was optimized.

References

- [1] Cody W, Kreulen J. T, Krishna V, Spangler S W, (2002) "The Integration of Business Intelligence and Knowledge Management", *IBM Systems Journal*, Vol. 41, No. 4, pp. 697-713, 2002.
- [2] Hong Yu and Vasileios Hatzivassiloglou. "Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences". In Michael Collins and Mark Steedman, editors, Proceedings of EMNLP-03, 8th Conference on Empirical Methods in Natural Language Processing, pages 129–136, 2003.
- [3] D. Houser and J. Wooders. Reputation in Auctions: Theory and Evidence from eBay. *Journal of Economics and Management Strategy*, 15:353-369, 2006.
- [4] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79–86, 2002.
- [5] A. B. Goldberg and X. Zhu, "Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization," in TextGraphs: HLT/NAACL Workshop on Graph-based Algorithms for Natural Language Processing, 2006.
- [6] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in Proceedings of WWW, pp. 519–528, 2003.
- [7] Ye, W. Shi and Y. Li. 2006. Sentiment classification for movie reviews in Chinese by improved semantic oriented approach. In Proceedings of 39th Hawaii International Conference on System Sciences, 2006.
- [8] L. Zhuang, F. Jing, X.-Y. Zhu, and L. Zhang, "Movie review mining and summarization," in Proceedings of the ACM SIGIR Conference on Information and Knowledge Management (CIKM), 2006.
- [9] Peter Turney. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the Association for Computational Linguistics (ACL)*, pages 417–424, 2002.
- [10] D.E. Rumelhart, G.E. Hinton and R.J. Willams, "Learning Representations by BackPropagating Errors," *Nature* 323, pp. 533-536 (1986).
- [11] J. L. Elman. Finding structure in time. *Cognitive Science*, 14:179{211, 1990.

Image Registration: An Application of Image Processing

Neeraj Kumar Pandey¹, Ashish Gupta², Amit Kumar Mishra³, Sanjeev Sharma⁴

**(Department of Computer science, Bhagwant Institute of Technology Muzaffarnagar U.P., India)*

*** (Department of Computer Science, Dev Bhoomi Institute of Technology Dehradun Uttarakhand, India)*

**** (Department of Computer science, Uttaranchal Institute of Technology dehradun Uttarakhand, India)*

***** (Department of Computer science, Uttaranchal Institute of Technology dehradun Uttarakhand, India)*

Abstract: Image registration is the process of overlaying one or more image to a reference image of the same scene taken at different time, from different view point and/or different sensor. Difference between images is introduced due to different imaging condition such that yields highest similarity between the input and the reference images. The objective of the registration process is to obtain the spatial transformation of an input image to a reference image by which similarity measure is optimized between the two images. There are a number of similarity measure is available which is used in the registration process. Out of which, a similarity measure which is based on information theory, called mutual information. Mutual information compare the statistical dependency between images. Registration based on mutual information is robust and could used for a large class of mono-modality and multimodality images. Image registration can be regarded as optimization problem where there is a goal to maximize the similarity measure. In this work we use mutual information as the similarity measure. There is a requirement to finding the global maxima of similarity measure. In this work we use simple genetic algorithm, share genetic algorithm, genetic algorithm combined with hill climbing algorithm for optimization. Being met heuristic these optimization technique require several decision to made during implementation, such as encoding, selection method and evolution operator. In this work we use two selection method roulette-wheel method and tournament selection method. Result indicates that these optimization techniques can be used for efficient image registration.

Keywords: computer tomography (CT), digital elevation models (DEM), Genetic algorithm (GA), Geographic Information System (GIS), mutual information (MI), magnetic resonance image (MRI), magnetic resonance spectroscopy (MRS), nuclear magnetic resonance (NMR), positron emission tomography (PET), single photon emission computed tomography (SPECT).

I. INTRODUCTION

Registration is the determination of a geometrical transformation that aligns Points in one view of an object with corresponding points in another view of that Object or another object. We use the term “view” generically to include a three-dimensional image, a two-dimensional image, or the physical arrangement of an Object in space. Difference between images is introduced due to different imaging condition such that yields highest similarity between the input and the reference images. Image registration geometrically aligns two

images the reference image and input image. Image registration is a crucial step in all image analysis tasks in which the final information is gained from the combination of various data sources like in image fusion, change detection, and multichannel image restoration. Typically, registration is required in remote sensing(multispectral classification, environmental monitoring, change detection, image mosaicing, weather forecasting, creating super-resolution images, integrating information into geographic information systems (GIS), in medicine(combining computer tomography (CT) and NMR data to obtain more complete information about the patient, monitoring tumor growth, treatment verification, comparison of the patient’s data with anatomical atlases),in cartography (map updating), and in computer vision (target localization, automatic quality control), to name a few.

In general, its applications can be divided into four main groups according to the manner of the image acquisition:

Different viewpoints (multi view analysis):

Images of the same scene are acquired from different viewpoints. The aim is to gain larger a 2D view or a 3D representation of the scanned scene. Examples of applications: Remote sensing mosaicing of images of the surveyed area. Computer vision shape recovery (shape from stereo).

Different times (multi temporal analysis):

Images of the same scene are acquired at different times, often on regular basis, and possibly under different conditions. The aim is to find and evaluate changes in the scene which appeared between the consecutive images acquisitions. Examples of applications: Remote sensing—monitoring of global land usage, landscape planning, Computer vision automatic change, detection for security monitoring, motion tracking, Medical imaging monitoring of the healing therapy, monitoring of the tumor evolution.

Different sensors (multimodal analysis):

Images of the same scene are acquired by different sensors. The aim is to integrate the information obtained from different source streams to gain more complex and detailed scene representation. Examples of applications are Remote sensing—fusion of information from sensors with different characteristics like panchromatic images, offering better spatial resolution, color/multispectral images with better spectral resolution, or radar images independent of cloud

cover and solar illumination. Medical imaging—combination of sensors recording the anatomical body structure like magnetic resonance image (MRI), ultrasound or CT with sensors monitoring functional and metabolic body activities like positron emission tomography (PET), single photon emission computed tomography (SPECT) or magnetic resonance spectroscopy (MRS). Results can be applied, for instance, in radiotherapy and nuclear medicine.

Scene to model registration:

Images of a scene and a model of the scene are registered. The model can be a computer representation of the scene, for instance maps or digital elevation models (DEM) in GIS, another scene with similar content (another patient), ‘average’ specimen, etc. The aim is to localize the acquired image in the scene/model and/or to compare them. Examples of applications are Remote sensing-registration of aerial or satellite data into maps or other GIS layers.

Computer vision targets in template matching with real-time images, automatic quality inspection and Medical imaging. Comparison of the patient’s image with digital anatomical atlases, specimen classification is also an application of image restoration.

1.1 Proposed work

The paper is organized as follows. The second section is a literature review of the area that gives the field background on the image registration process. The third section reviews the optimization methods of concerning image registration. The fourth section focuses on methodology and implementation issues. Results from the investigation are presented in the fifth section. Finally the paper is concluded with future work in the area.

II. IMAGE REGISTRATION PROCESS

The registration process involves finding a single transformation imposed on the input image by which it can align with the reference image. It can be viewed as different combination of choice for the following four components. [12].

- (1) Feature space
- (2) Search space
- (3) Similarity measure
- (4) Search strategy

The **Feature space** extracts the information in the images that will be used for matching. The **Search space** is the class of transformation that is capable of aligning the images. The **Similarity measure** gives an indication of the similarity between two compared image regions. The **Search strategy** decides how to choose the next transformation from the search space, to be tested in the search to spatial transformation. This work focuses on image registration of two medical images of having different modality i.e image acquired with different sensor e.g. images, MRI images. We consider set of image pixel intensity as the feature space and affine transformation as the search space. A popular similarity

measure is mutual information (MI) consider as the mutual information .MI is based on the information theory. MI compares the statistical dependency between images. Registration based on the MI is robust and can be used for a large class of images acquired by the same sensor and different sensors. For the search strategy we use simple genetic algorithm and share genetic algorithm. Genetic algorithm (GA) is based on the concept of the natural process of specie evolution to realize simple and robust methods for optimization. GA is a stochastic technique for optimization, convergence towards global optima is very slow. To improve the time constraint of the registration process we apply simple genetic algorithm combined with the hill climbing algorithm. Hill climbing algorithm is a local search algorithm and execution is fast. In this work, we perform a comparative study of the image registration process on the multimodal medial images by using different genetic algorithm relative to the performance as accuracy and time. We use two genetic algorithm as simple genetic algorithm , sharing genetic algorithm using two selection criteria as roulette-wheel selection and tournament selection.

Probability distribution of gray values can be estimated by counting the number of times each gray value occurs in the image and dividing those numbers by the total number of occurrences. An image consisting of almost a single intensity will have a low entropy value; it contains very little information. A high entropy value will be yielded by an image with more or less equal quantities of many different intensities, which is an image containing a lot of information.

In this manner, the Shannon entropy is also a measure of dispersion of a probability distribution. A distribution with a single sharp peak corresponds to a low entropy value, whereas a dispersed distribution yields a high entropy value. Summarizing, entropy has three interpretations: the amount of information an event (message, gray value of a point) gives when it takes place, the uncertainty about the outcome of an event and the dispersion of the probabilities with which the events take place

III. OPTIMIZATION THEORY

This section reviews the theory behind the optimization methods Genetic Algorithm, Share Genetic Algorithms and Hill climbing Algorithm. Some research regarding the implementation of the methods on image registration is reviewed.

3.1 Search Space

When solving an image registration problem, we look for a particular solution that will be better than all or almost all other feasible solutions. Depending on the number of parameters n that constitute a solution, an n -dimensional *search space* consisting of the set of all possible solutions is created. If we mark each point in the search space with the corresponding cost for that solution, we get a landscape-like hyper surface. Our aim with the search is to find the lowest valley in this landscape. This is often a rather time consuming process, since the hyper surface rarely behaves in a smooth and predictable way.

3.2 Genetic Algorithms

This section describes the *Genetic Algorithms* [7] (GA), a random optimization technique inspired by the theory of evolution and the survival of the fittest. Genetic algorithms belong to the broad class of *Evolutionary Algorithms* (EAs) which take inspiration from nature's own way of evolving species.

3.2.1 Background

The idea of *evolutionary computing* was developed in the 1960's and has since then developed into a significant area within Artificial Intelligence. We will focus particularly in the concept of *genetic algorithms*, invented and developed by John Holland. The problem solving methodology employed in a genetic algorithm closely resembles an evolutionary process, where successively more and more fit solutions are "evolved" through an iterative procedure.

3.2.2 Algorithm Description

The operations of the genetic algorithm are very simple. It maintains a population $x_1 \dots x_n = \{x_1, \dots, x_n\}$ of n individuals x_i (which may consist of a vector of parameters). These individuals are candidate solutions to some objective function $F(x_i)$ that is to be optimized to solve the given problem. The individuals are represented in the form of 'chromosomes,' which are strings defined over some alphabet that encode the properties of the individuals. More formally, using an alphabet $A = \{0, 1, \dots, k-1\}$, we define a chromosome $C = \{c_1, \dots, c_l\}$ of length l as a member of the set $S = A^k$, i.e., chromosomes are strings of l symbols from A . Each position of the chromosome is called a *gene*, the value of a gene is called an *allele*, the chromosomal encoding of a solution is called the *genotype*, and the encoded properties themselves are called the *phenotype* of the individual. In the GA, typically a binary encoding is used, i.e., the alphabet is $A = \{0, 1\}$.

IV. ALGORITHM DESCRIPTION

Being meta heuristic GA require several decision to be made during implementation for encoding, selection, crossover and mutation.

Encoding

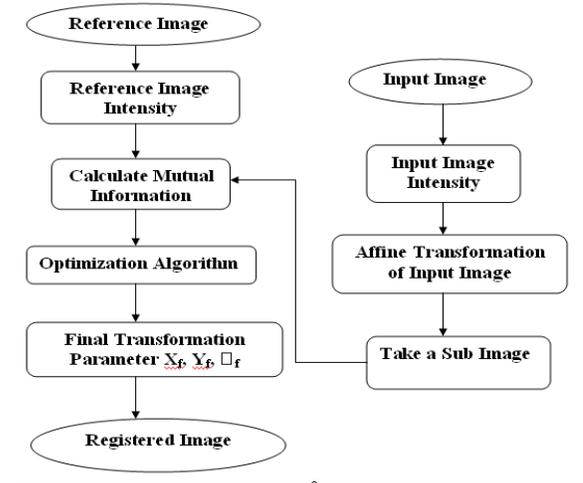
The first decision to take when implementing a GA is how solution states should be encoded into chromosomes. Some encoding techniques are

- Binary Encoding
- Octal Encoding
- Hexadecimal Encoding
- Gray Encoding
- Floating Point Encoding

V. METHODOLOGY

As we know that image registration is the process of overlaying one or more images to reference of the same seen. The flow graph for the registration process shown in fig below. There are four main step such as (i) feature space (ii) search space (iii) search strategy and (iv) similarity measure

in the image registration process. The experiment is done in MATLAB 7.5. The registration process is implemented for the multimodal images (image of different sensor). The implementation aspects of these steps are as follows.



Flow Graph of Image Registration using mutual information

Figure 1

V. RESULTS AND DISCUSSION

We test the image registration of the 7pair of medical images using the following algorithm.

- Simple Genetic Algorithm using roulette-wheel selection (GAr)
- Shared Genetic algorithm using roulette-wheel selection (SGAr)
- Genetic Algorithm combined with Hill-climbing algorithm using roulette-wheel selection (GAr+Hill)
- Simple Genetic Algorithm using tournament selection (GAt)
- Shared Genetic algorithm using tournament selection (SGAt)

VI. CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this work, we implement two genetic algorithms with two selection criteria i.e. simple genetic algorithm and share genetic algorithm. We also implement a hybrid algorithm genetic algorithm combined with hill climbing algorithm. We conclude from the extracted result as follows

- All the three algorithms, simple genetic algorithm, share genetic algorithm, and genetic algorithm combined with hill climbing algorithm are feasible alternative in performing image registration.
- Genetic algorithm represents an effective technique in multimodal optimization problem. One problem with genetic algorithm is that it can be trapped in local minimum due to genetic drift. By sharing genetic algorithm the problem can be solved. So that sharing genetic algorithm has given better performance than the genetic algorithm. But the algorithm is highly sensitive to

calibration parameter. Therefore some time it does not give better performance.

- (iii) The proposed algorithm genetic algorithm combined with hill climbing algorithm give very good performance with respect time as well as accuracy. The genetic algorithm is time consuming which is over come by using the hill climbing algorithm. The drawback of hill climbing algorithm is that it gives local optima, which is overcome by using GA.
- (iv) In all the case the hybrid GA (GA+Hill) will give better performance in terms of accuracy to the simple GA and in some case it give better result than all the rest algorithm. In term of time it is fast among all the algorithm.
- (v) We use genetic algorithm and hill climbing algorithm sequentially. First the genetic algorithm executes after that the hill climbing algorithm where genetic algorithm gives near to option. As hill climbing algorithm is local search algorithm this may occur that the maxima which is achieved in genetic.

6.2 Future Work

Some of the future work possible in the area is listed as follows.

- (i) In this work we use mutual information as the similarity measure, other similarity measure such as gradient coded MI, weighted MI can also be used.
- (ii) Parallel GA can also be used, to improve performance.
- (iii) We use chromosome vector of four parameter translation along the x-axis and y-axis, rotation, scaling. The number of parameter can also increased to six parameter by taking two other parameter skewing and squeezing.
- (iv) Adaptive GA and other hybrid GA can be used to improve the time and achieve sub pixel accuracy.
- (v) Crowding GA can also be used to overcome the problem of genetic drift arises in the simple GA in multimodal optimization problem.
- (vi) We use hill climbing algorithm with GA sequentially, i.e first we use GA then hill climbing. It can also implement such that the hill climbing algorithm used within the GA.

ACKNOWLEDGEMENT

We would like to thank Prof. Yudhveer Singh Moudgil from Uttaranchal Institute of Technology, Dehradun for his continuous support and guidance throughout reviewing this research paper. His area of specialization is Cryptography & network Security, Digital Image Processing and Advance Computer Architecture. He is author of many technical books and published a plenty of research papers in national and international journals.

REFERENCES

- [1] Barbara Zitova*, Jan Flusser, "Image registration methods: a survey", *Image and Vision Computing* 21 (2003) 977–1000.
- [2] Mark P. Wachowiak, Renata Smolíková, Yufeng Zheng, Jacek M. Zurada, and Adel S. Elmaghraby, "An Approach to Multimodal Biomedical Image Registration Utilizing Particle Swarm Optimization", *IEEE*

TRANSACTIONS ON EVOLUTIONARY COMPUTATION, VOL. 8, NO. 3, JUNE 2004.

- [3] Geoffrey Egnal, Kostas Daniilidis, "Image Registration Using Mutual Information" University of Pennsylvania Department of Computer and Information Science Technical, Report No. MS-CIS-00-05.
- [4] Frederic Maes, Dirk Vandermeulen, and Paul Suetens, "Medical Image Registration Using Mutual Information", *PROCEEDINGS OF THE IEEE*, VOL. 91, NO. 10, OCTOBER 2003.
- [5] Giuseppe Pascal, Luigi Troiano, "A Niche Based Genetic Algorithm For Image Registration", *Proceedings of the ICEIS 2007 - Ninth International conference on Enterprise Information Systems*, Volume AIDSS, Funchal, Madeira, Portugal, June 12-16, 2007.
- [6] Josien P. W. Pluim, J. B. Antoine Maintz, and Max A.iergever. "Mutual-Information-Based Registration of Medical Images: A Survey", *IEEE TRANSACTIONS ON MEDICAL IMAGING*, VOL. 22, NO. 8, AUGUST 2003 .
- [7] Flávio Luiz Seixas, Luiz Satoru Ochi, Aura Conci, Débora C. M. Saade, "Image Registration Using Genetic Algorithms", *GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA*. ACM 978-1-60558-130-9/08/07.
- [8] Jharna Majumdar and Y. Dilip, "Implementation of Image Registration Algorithms for Real-time Target Tracking Through Video Sequences", *Defence Science Journal*, Vol. 52, No. 3, July 2002, pp. 227-242
- [9] Maslov, I. V. and Gertner, I. (2001). Gradient-based geneticalgorithms in image registration. In *Proc. SPIE, AutomaticTarget Recognition XI: AeroSense 2001*, volume4379, pages 15–34, Orlando, FL, USA.
- [10] Dasgupta, D. and McGregor, D. R. (1992). Digital imageregistration using structured genetic algorithms. In *Proceedings of SPIE the International Society for OpticalEngineering, Vol. 1766*, pages 226–234.

A Neural Network Model of Face Detection for Active Vision Implementation

Yasuomi D. Sato^{*,**}, Yasutaka Kuriya^{*}

^{*} Department of Brain Science and Engineering, Graduate School for Life Science and Systems Engineering, Kyushu Institute of Technology, Japan

^{**} Frankfurt Institute for Advanced Studies (FIAS), Goethe University Frankfurt, Germany

Abstract : We suggest an active vision design in a neural style to achieve face detection, or face recognition in gaze, understanding network structures on impaired face recognition in developmental disorders such as autism spectrum deficits and Asperger's syndrome. The core of the active vision design is so-called corollary discharge (CD) on an ascending pathway of the superior colliculus to the frontal eye field (FEF) via the mediodorsal thalamus. The CD is the signal about upcoming eye movements for achieving the stable visual perception. Assuming that the CD is parameterized as controlling dynamical projection creation between the FEF and the Fusiform gyrus, face detection performance is shown to gradually decrease when the value of the CD parameter increases. This indicates a deficit of accurate eye motion, which may elicits impaired face recognition in gaze.

Keywords – Active vision design, Face detection, Corollary discharge, Dynamic link architecture, Elastic graph matching

I. INTRODUCTION

The deficits of face recognition (or face perception) are frequently diagnosed in developmental disorders such as autism spectral disorders (ASDs) and Asperger's syndrome symptoms [1]. They are seen even in schizophrenia patients [2]. The face recognition deficiencies evoke serious social problems on developmental process in the children, because face recognition is one of very significant requirements to acquire social communication skills in childhoods. The urgent task is thus to correctly understand and then solve a mechanism on impaired face recognition. This will be allowed to remove social problems occurred in the patients and enable them to healthy come back to common life.

Fusiform Gyrus (FFG) is regarded as a key brain area in face recognition. A deficit in cholinergic innervations of the FFG was observed in adults with ASD. The deficit may be related to not only current but also childhood impairment of social functioning [3]. But why the cholinergic activity is reduced is still unclear, although the cholinergic activity has been well-known to regulate the function of the visual pathway, including the FFG. To answer this question, we will have to overview a whole configuration of visual pathway, including a visuomotor system and so forth. In addition, we will have to study cognitive modeling of visual perception,

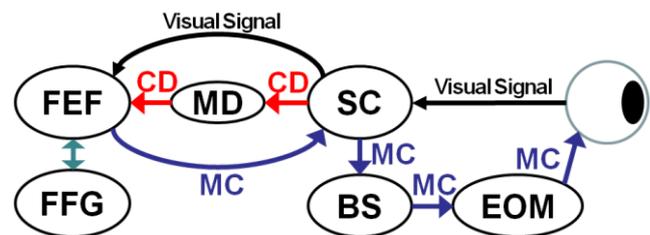


Fig. 1 Two pathways of corollary discharge (CD) and motor command (MC) after visual signals are transmitted to the superior colliculus (SC). The ascending pathway of the CD: The SC to the Frontal eye field (FEF) via the mediodorsal thalamus (MD), which projects to the Fusiform gyrus (FFG). The descending pathway of the MC: The SC to the extraocular muscles (EOM) via the Brain Stem (BS). This figure is referred to [4-6].

In general, an object always and complicatedly moves. The bulks of eye also move such as saccade. Nevertheless, visual perception for moving objects keeps stably reflected in the brain, possibly, in the frontal eye field (FEF)). Such visual stabilities are not accomplished with impaired corollary discharge (CD). In Fig. 1, the CD pathway of the superior colliculus (SC) to the FEF through the mediodorsal thalamus (MD) was empirically confirmed to convey signals about forthcoming eye movements (that is, when a saccade will happen and where it will go) [4]. Such an ascending signal of the CD allows the FEF to feedback motor command signals (called MC in Fig.1) to the SC, to generate a saccade for precise object tracking.

However [5] shows that inactivation in the MD can impair the CD, which makes it very difficult to keep track of an object. Therefore, this physiological experiment indicates that the CD plays an important role in achieving precise and accurate eye track of an object as well.

This is the same of achieving face recognition, keeping track of the face. If the CD can contribute to face recognition in gaze, it leads to understanding deficits of the face tracking (or face attention) by impaired CD pathway. In an fMRI experiment, an intensive connectivity between the FEF and FFG is already observed [6]. If a mechanism on visual face recognition on a network of the FEF and FFG with the CD pathway is clear, it can be expected as helpful to understand mechanisms on face recognition deficits in developmental disorders. For this purpose, it is very important to study a neural model of recognition/detection of a face, including MC signals on CD pathway.

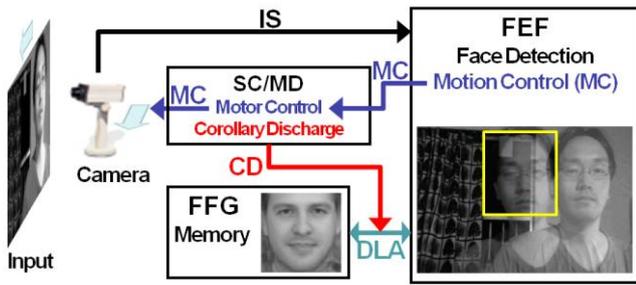


Fig. 2. An active vision configuration for face recognition. Image signals (ISs) are delivered to the FEF from a camera, in which a motion of the face is calculated when detecting a position of the face on an input image. Information about the motion as a motor command (MC) signal descends to the camera via the SC/MD module. It thus allows the camera to smoothly track a face. In the SC/MD module, a parameter of the corollary discharge (CD) controls matching process for face recognition/detection in the dynamic link architecture (DLA). Here shorten forms of the SC/MD, FEF and FFG is referred as in Fig.1.

We propose an active vision design in a neural style for achieving face detection. The dynamic link architecture (DLA) [7-9] employed in this work can frequently be regarded as a powerful tool for explaining two following points: (1) How neural activities in the FEF can be projected to the ones in the FFG, including the ascending pathway of the CD. (2) How face recognition (or detection) can be achieved. Assuming that the CD plays an important role in dynamical creation of projections of the FEF to FFG, we show face detection performance increased with changes of the CD as a control parameter. Within a framework of the active vision design, we will discuss considerable mechanisms on face recognition deficits observed in patients with developmental disorders.

II. ACTIVE VISION DESIGN

2.1. System Configuration

Let us begin by looking at a whole system for achieving face recognition in gaze. This mainly consists of three

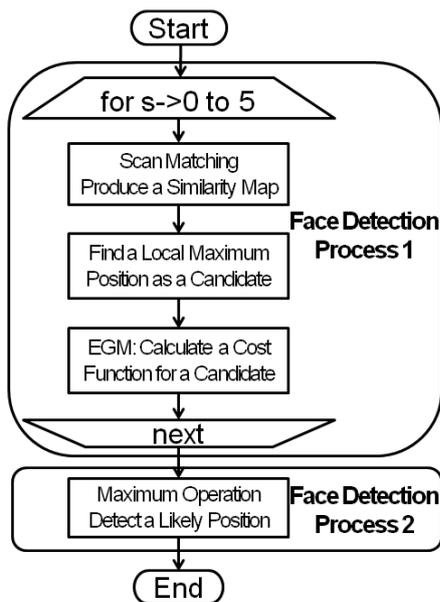


Fig. 3. A flowchart of face detection process. It consists of two sub-processes. The first (Face Detection Process 1) is candidate position finding while the second (Face Detection Process 2) is the most likely position detection.

modules, for the simplicity, the FFG, the FEF, as well as the SC/MD (see Fig.2). The FFG as memory stores a facial image (called the M). In the FEF, a face is detected on an input (I) image, which is sent as an image signal (IS) by a camera. The face detection is done by matching the images I to M within a framework of the DLA. The matching process is controlled by the CD parameter in the SC/MD. Motion is estimated in comparison with the previous frame. It is delivered as a motor command (MC) signal to a servomotor in the camera through the SC/MD. Thus, the camera can track a face.

2.2. Feature Representation in FFG module

We employ an image of the average face created by a face generator in [10], whose size is $A_0=60 \times 60$ pixels. Since feature representations are composed of the different scales, the image is down-sampled with a_0^s ($s=0, \dots, 5$), called the M_s . Here $a_0=0.85$ and $s=0$ implies the original size of the image M . A square graph of $(n \times n - 4)$ nodes without any vertexes is set on each resolution image M_s . Here n ($=5$) is the number of full nodes on a row and column of the square graph. Each node on the image M_s is convoluted with a family of Gabor functions $\Psi_r(z)$. r ($=0, \dots, 7$) is an orientation parameter. The Gabor feature usually consists of 8 different orientation components. Each orientation component is given by a convoluted value ($\hat{J}_r^{M_s}$):

$$\hat{J}_r^{M_s}(p_s) = \int F^{M_s}(p_s - \hat{p}_s) \Psi_r(p_s - \hat{p}_s) d^2 \hat{p}_s, \quad (1)$$

$$\Psi_r(p_s) = \frac{k_r^2}{\sigma^2} \exp\left(-\frac{k_r^2 p_s^2}{2\sigma^2}\right) \left[\exp(ik_r p_s) - \exp\left(-\frac{\sigma^2}{2}\right) \right]. \quad (2)$$

$\sigma=2\pi$. The wave number vector k_r can be expressed as

$$\vec{k}_r = \begin{pmatrix} k_{rx} \\ k_{ry} \end{pmatrix} = \begin{pmatrix} k \cos \varphi_r \\ k \sin \varphi_r \end{pmatrix}, \quad k = \frac{\pi}{2}, \quad \varphi = \frac{\pi}{8} r. \quad (3)$$

One orientation component $J_r^{M_s}$ in the Gabor feature takes an absolute value of $\hat{J}_r^{M_s}$:

$$J_r^{M_s}(p_s) = |\hat{J}_r^{M_s}(p_s)|. \quad (4)$$

We mention physiological backgrounds of multi-scale feature representation for a face. Physiological experiments for receptive field already finds that neurons being tuned to high spatial frequencies have narrower tuning range than neurons being tuned to low spatial frequencies. Also, the receptive field structure represented by Eq. (2) is observed as constructed by multiplying a global sinusoidal grating by a bell-shaped Gaussian envelope [11].

Furthermore, human vision can be conceived to be achieved through low pass filter processing [12]. Once an input image about some environmental scene is received on a retina, the highest spatial frequency component of the Gabor filter is sequentially discarded in bottom-up flow (referred to [13]). The prospective discarding spatial frequency elements may be stored in another area through another pass in the visual cortex. The FFG can be considered as one of candidates. Therefore, the multi-scale Gabor feature representations are suitable for modeling of memory.

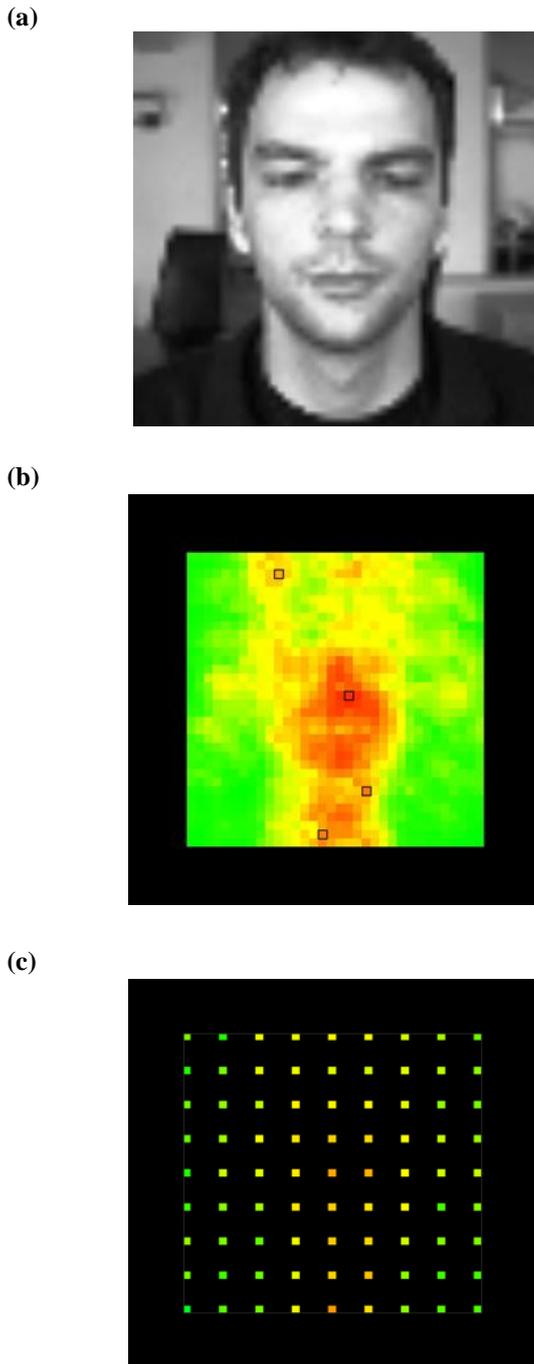


Fig. 4. Similarity maps when a static grayscale image of a face, used as an input image (a), is scanned with a square graph at M_s . (b) $n^{cd}=1$ and (c) $n^{cd}=4$ where n^{cd} is the scanning step number. In (b) and (c), black squares show a local maximum position on the similarity map.

2.3. CD controlled Graph Matching in FEF module

In analogy, each frame on the input (I) image is convoluted with a family of Gabor function. The main face detection processing is then proceeded. As shown in Fig.3, the face detection processing consists of two sub-processes of candidate position finding and the most likely position detection.

2.3.1. Scan Matching

In the first sub-process, in order to pick up candidate positions, an entire or fragmentary similarity map for each scale M_s is calculated with scan matching of the undistorted graph G_s onto the image I :

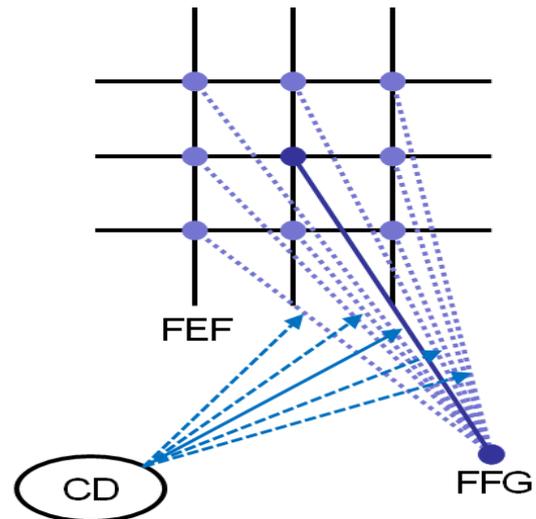


Fig. 5. Dynamic link architecture (DLA) between the FEF and FFG. One feature detector for model representation on the FFG, which has been projected to all of feature points on the FEF, then finds the optimal point that takes the highest similarity to the model representation through the CD control (see a solid line and a solid arrow). In the CD, the other controllers (broken arrows) weaken projections between the FEF and the FFG (broken lines).

$$e_s^S(x_l) = \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\sum_r J_r^I(x_l) \cdot J_r^{M_s}(p_s)}{\sqrt{\sum_r (J_r^I(x_l))^2 \sum_r (J_r^{M_s}(p_s))^2}}, \quad (5)$$

where $e_s^S(x_l)$ is the similarity of the Gabor feature at x_l on the image I to the one for each node p_s . It is noticed that G_s sometimes represents a set of nodes of the graph. In the scan matching process, the left-upper of the image M_s on which a square graph is being set up, is firstly adjusted to the left-upper of the I . The M_s is repeatedly scanned to the right for each row on the I . The scan ends when the right-lower of the M_s arrives at the right-lower of the I (not shown here). As a result, we obtain a similarity map at each image M_s . We exemplify similarity map results when one facial image I is scanned with a square graph at the level M_s (Fig. 4).

Next, we pick up some candidate positions that can be expected as the one of a face. Here let the candidate position x_l^c be defined as the center of the image M_s or of its square graph. It corresponds to the local maximum on the similarity map. The local maximum is satisfied with a gradient condition when all differences of the x_l^c to its nearest neighbors x_l^n take positive:

$$e_s^S(x_l^c) - e_s^S(x_l^n) > 0. \quad (6)$$

The candidate pixel is depicted with a black square on the similarity map as shown in Fig. 4(b).

2.3.2. Elastic Graph Matching (EGM)

The EGM for each candidate position x_l^c is computed to obtain the maximum value of the cost function E_s^c , which is given by

$$E_s^c(x_l) = e_s^S(x_l) - \lambda_d e_s^d(x_l), \quad (7)$$



Fig.6. Facial images of a same or different person with the background in a Bio ID dataset of 1521 images.

$$e_s^d(x_I) = \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\left| \sum_{p'_s \in G'_s} D_{p_s p'_s}^{M_s} - \sum_{p'_s \in G'_s} D_{x_I p'_s}^I \right|}{\left| \sum_{p'_s \in G'_s} D_{p_s p'_s}^{M_s} \right|}, \quad (8)$$

$$+ 1 - \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\left| \sum_{p'_s \in G'_s} A_{p_s p'_s}^{M_s} \cdot \sum_{p'_s \in G'_s} A_{x_I p'_s}^I \right|}{\left| \sum_{p'_s \in G'_s} A_{p_s p'_s}^{M_s} \right| \cdot \left| \sum_{p'_s \in G'_s} A_{x_I p'_s}^I \right|}$$

where $e_s^d(x_I)$ represents the elasticity of the graph on the image I . λ_d is a constant parameter for the graph elasticity. $\lambda_d=0.05$, except for obtaining a similarity map $E_s(x_I)$ when $\lambda_d=0$. G'_s is a set of nearest neighbor nodes p'_s for p_s . $D_{p_s p'_s}^{M_s}$ and $D_{x_I p'_s}^I$ are the Euclidean distance between nodes p_s (or x_I) and p'_s on the graph of the image M_s (or I). $A_{p_s p'_s}^{M_s}$ and $A_{x_I p'_s}^I$ take one vector form consisting of 4 elements. Each element is an angular between two nearest neighbors on each quadrant, centered at p_s .

Each node on the image I , which corresponds to the node p_s^c of the square graph M_s , surveys an optimal pixel p_s^I taking a maximum of the cost function E_s^c within a search region R :

$$p_s^I = \max_{x_I \in R} \{E_s^c(x_I)\} \quad (9)$$

where R is a set of pixels that can be picked up in a square with the size of $(2q+1) \times (2q+1)$, centered at the pixel corresponding to p_s^c . $q=4$.

The optimal pixel x_m^I is singled out with the maximum operation of all candidates, which must be the most invariant to feature representation for an M face:

$$x_m^I = \max_{c \in C} \{E_s^c(p_s^I)\} \quad (10)$$

Here C is a set of candidates.

2.4. Motor Command

Finally, let us explain briefly an algorithm of face tracking by controlling a camera. In this algorithm, motions of a face are defined as a difference of the face position on the $I^{(i)}$ to



Fig.7. Face detection demonstration when we used a camera. Numerals on the left-upper are respectively the maximum value of a cost function and computational cost value for each frame.

the one on $I^{(i+1)}$

$$x_m^{I^{(i+1)}} - x_m^{I^{(i)}} \quad (11)$$

where i th is the frame index. The information about motions of a face is transmitted to a motor controller in the servo camera via the SC/MD module. The camera can be expected to track a face on the input.

A goal of this article is to propose active vision architecture to understand a mechanism on impaired face detection. It is naturally expected that the accuracy of motion gets worse when face detection performance decreases. Therefore, we omit the concrete explanation about the motor control algorithm in this article.

2.5. Modeling of Corollary Discharge

A role of the corollary discharge (CD) pathway is reconsidered for interpreting the face detection system in a neural style. As mentioned in the Introduction, deficits of the CD decrease the accuracy of eye movements. The CD is thus necessary to control precisely motions of the eye. To increase the accuracy of motion, the accuracy of the face detection also increases.

For this, the core of face detection with scan matching and EGM is feature-based correspondence finding within a framework of the DLA. As shown in Fig.5, one feature detector on the M graph, projected to all pixels in the I , tries to find the optimal with topographic mappings. It also enables us to find the highest invariance to feature representations for a face. This is the essential of searching the local maximum of a similarity map. Also, the same is finding the maximum of a cost function.

However there is still a problem in searching the local maximum, which detection fault is increased if a number of the local maximum is found as shown in Fig. 4(b). As one possibility to solve this problem, we do not create a whole similarity map, but the local similarity map as shown in Fig. 4(c). This can be expected to improve the face detection fault.

Return to neural model interpretations of our face detection system. The scanning step number n^{cd} controls projection patterns. The all-to-all projection pattern is, for

TABLE I
FACE DETECTION PERFORMANCE

Scan Step	The number of correct detected images	The detection rate (%)
$n^{cd}=1$	1485	97.6
$n^{cd}=2$	1474	96.9
$n^{cd}=3$	1472	96.7
$n^{cd}=4$	1445	95.0

A face detection ability test using the database of the BioID with 1521 facial images. An undistorted graph is scanned with scan step n^{cd} on the input image.

example, changed to the sparse when the n^{cd} increases. It addresses that the CD controls neural connectivity patterns between the FEF and FFG. The CD control parameter, namely, the scanning step number is one of the remarkable points in our detection system. The face detection rate computed with our detection system is related to the accuracy of the face detection. This will thoroughly be discussed in the next section.

III. DETECTION PERFORMANCE

We test an ability of our face detection system, using the Bio ID database [14] that involves 1521 facial images. Fig. 6 shows some of facial images in the Bio ID database. On the other hand, an image of an average face produced with many German facial photos in [10] is employed as the model (M) image. This indicates that our face detection system has recognition ability for a face as the object, not personal identity.

In the detection ability test, we do 4 trials of $n^{cd}=1, 2, 3$ and 4. For each trial, the M image should be tried to match to each facial image in the Bio ID database. Correct face detection can be defined when eyes are in 3 of 5th from the top of detected square area. The correct face detection rate is the accuracy of the face detection in this work. The correct detection results are shown in TABLE I.

In TABLE I, when the n^{cd} is increased 1 to 4, the face detection rate is gradually decreased 97.6% to 95.0%. When the scan step number is furthermore increased, the face detection rate becomes decreased (as not shown here), because the number of candidates is decreased.

These results indicate as follows: The first can expect more improvement of face tracking in active vision. In fact, we use a camera. Our face detector can, in real time, capture a face without any severe problems even though the size of the face is bigger or smaller (see Fig. 7). The second is the validity of neural model interpretation of the corollary discharge. Increases of the scan step represent because our model supports experimental results for a functional role of the corollary discharge.

IV. DISCUSSION

We propose active vision architecture in a neural style, assuming that this architecture contains a mechanism to control a motor in a servo camera. The key is the bottom-up signal of the MD to the FEF for information about forthcoming eye movement, which is called the CD, on which the top-down signal is transmitted as a motor

command. The functional role of the CD is to improve or preserve eye movement accuracy.

In this article, making full use of the concept of the DLA, we parameterize the role of the CD for its simplicity, and also show that face detection performance calculated here, which qualitatively supports to an experimental result, is decreased with an increase of the CD parameter value. It is not doubtful that the CD is a necessarily function to object tracking or visual attention.

It is still unclear if or not the impaired CD pathway is a central mechanism on deficits of face recognition. This is due to a decline of face detection performance indirectly means difficulties of face tracking by motor controls. Sect. I Introduction already reports that patients with deficits of face recognition can difficultly pay attention to, or track a face. It can thus be inferred that they are not in healthy communication environment in their childhood so that the CD pathway is impaired in the brain. This may cause a decline of abilities for attention to and track a face to induce deficits of face recognition.

As shown in Fig. 5(c), we have created a sparsely distributed map of the similarity. Since it means a decrease of correlations or connectivity intensities between neural activities of the FEF and FFG, the localized similarity map result may support inactivation of the cholinergic system in the FFG. However, to confirm support of the computational model to the cholinergic inactivation, we will have to study more declines of neural activities in FEF and MD, together with cholinergic inactivation in the FFG in the experiment. Such experiments are not yet reported as far as we know.

There are still lots of improvements in our face detection algorithm. The first priority is the CD deficit parameter n^{cd} . One of goals in this work is to understand functional mechanisms on achievement of face recognition in developmental process. We have to study neural network modeling of the related learning process. For this, physiological experiments that have been done by Bell et al. [15] are of great use as a reference.

Bell et al. addressed that representations of recent sensory input that followed motor commands are stored and updated through anti-Hebbian plasticity at synapses between corollary discharge conveying fibers and the sensory area. The updating is prevented by no motor commands and no plastic changes. If such a learning model is achieved for being implemented into our face detection system with the EGM, this can expect to lead to solving mechanisms on how face recognition can be impaired. Also, it gives us one of solutions that the deficit of the CD harms not only to the FEF activation, but also to updating and storing input image in the FEF. Correspondingly the deficit elicits the cholinergic inactivation in the FEF. Furthermore, the motor command signals can no longer be transited for accurate eye movements so that eye tracking or attention to a face is impossible. Therefore, modeling of the learning on the CD pathway is crucial for giving us explanations of symptoms in face recognition disorders.

Taking a glance at a current research field of developmental robotics, researches and developments on physiologically inspired active vision and its relevant learning algorithm does not yet seem to be reported [16]. The learning model implemented into cognitive

developmental robots [17] is the already existing. It cannot be in a neural style, much less modeling of a corollary discharge mechanism. In [18], a multiple forward model for a corollary discharge mechanism is used to computationally simulate experiments on attribution of own actions to intention of self or others. Such a corollary discharge model is still conceptual or cognitive, is not based on the physiological principles, is already done on computational simulations, and is not yet applied into robots.

There is ample scope for some progresses of computer vision technology in a face detection system we proposed as a neural style. Even in the improvement [19], our face detection system has higher face detection performance and comparably rapid computational speeds, compared to the Viola-Jones face detection algorithm [20]. The details are removed here, because they are different directions to this work. But, in the near future, real-time visual information processing models will be achieved on computers. They can be competed with real visual processing in the brain.

V. CONCLUSION

In this article, we studied a neural network model for correct face detection, to understand functional mechanisms on impaired face recognition in developmental disorders such as autism spectrum deficits and Asperger's syndrome. The core of the design is a so-called corollary discharge (CD) mechanism on an ascending pathway of the superior colliculus to the frontal eye field (FEF) via the mediodorsal thalamus. The CD is the signal about upcoming eye movements for achieving the stable visual perception. We assume that the CD plays an important role in dynamical creation of projection of the FEF to the Fusiform gyrus and then show that face detection performance decreases with changes of the CD parameter's value. This can predict a decline of the accuracy of eye motion. From a result of the detection performance decline, considerable mechanisms on the face recognition deficit were discussed, in terms of difficulties of face tracking. We indicated that an active vision design proposed here was neurally plausible as well as powerful to explain deficiencies of face recognition.

REFERENCES

- [1] M. L. Phillips, Facial processing deficits and social dysfunction: how are they related? *Brain*, 127(8), 2004, 1691-1692.
- [2] F. Irani, S. M. Platek, I. S. Panyavin, M. E. Calkins, C. Kohler, S. J. Siegel, M. Schachter, R. E. Gur, R. C. Gur, Self-face recognition and theory of mind in patients with schizophrenia and first-degree relatives. *Schizophrenia Research* 88, 2006, 151-160.
- [3] K. Suzuki, G. Sugihara, Y. Ouchi, K. Nakamura, M. Tsujii, M. Futatsubashi, Y. Iwata, K. J. Tsuchiya, K. Matsumoto, K. Takebayashi, T. Wakuda, Y. Yoshihara, S. Suda, M. Kikuchi, N. Takei, T. Sugiyama, T. Irie, Reduced Acetylcholinesterase Activity in the Fusiform Gyrus in Adults With Autism Spectrum Disorders. *Arch. Gen. Psychiatry* 68(3), 2011, 306-313.
- [4] M. A. Sommer, R. H. Wurtz, What the Brain Stem Tells the Frontal Cortex. I. Oculomotor Signals Sent From Superior Colliculus to Frontal Eye Field Via Mediodorsal Thalamus. *J. Neurophysiol.* 91, 2004, 1381-1402.
- [5] M. A. Sommer, R. H. Wurtz, What the Brain Stem Tells the Frontal Cortex. II. Role of the SC-MD-FEF Pathway in Corollary Discharge. *J. Neurophysiol.* 91, 2004, 1403-1423.
- [6] J. Xaing, J. Chen, The study on Functional Connectivity between Frontal Eye Field and Other Brain Cortex in Visual Processing. *2010 Int. Conf. on Computational Aspects of Social Networks (CASoN)*, 2010 333-336.
- [7] C. von der Malsburg, The correlation theory of brain function. *Internal Report, Max-Planck-Institut für Biophysikalische Chemie*, Postfach 2841, D3400 Gottingen, FRG, 1981.
- [8] C. von der Malsburg, How are nervous structures organized? in Synergetics of the Brain, *Proc. Int. Symp. Synergetics*, E. Basar, H. Flohr, H. Haken, and A. Mandell, Eds. Berlin, Germany: Springer, 1983, 238-249.
- [9] C. von der Malsburg, Nervous structures with dynamical links. *Berichte der Bunsengesellschaft für Physikalische Chemie* 89, 1985, 703-710.
- [10] <http://www.faceoftomorrow.com/>
- [11] M. Čadik, Human Perception and Computer Graphics. Czech Technical University Postgraduate Study Report, 2004.
- [12] N. Thibos, Image processing by the human eye. In: Visual Communications and Image Processing IV, William A. Pearlman (ed), *Proc. SPIE 1199*, 1989, 1148-1153.
- [13] Y. D. Sato, Y. Kuriya, Visual Constructed Representation for Object Recognition and Detection. In B.-L. Lu, L. Zhang, and J. Kwok (Eds.): *ICONIP 2011, Part III, LNCS 7094*, 611-620, Springer-Verlag Berlin Heidelberg, 2011.
- [14] O. Jesorsky, K. Kirchberg, R. Frischholz: Robust Face Detection Using the Haus-dorff Distance. In J. Bigun and F. Smeraldi editors, *Audio and Video based Person Authentication - AVBPA 2001*, Springer, 2001, 90-95.
- [15] C. C. Bell, A. Caputi, K. Grant, J. Serrier, Storage of a sensory pattern by anti-Hebbian synaptic plasticity in an electric fish. *Proc. Natl. Acad. Sci. USA* 90, 1993, 4650-4654.
- [16] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, C. Yoshida, Cognitive Developmental Robotics: A Survey. *IEEE Transactions on Autonomous Mental Developmental* 1(1), 2009, 12-34.
- [17] A. Watanabe, M. Ogino, M. Asada, Mapping facial expression to internal states based on intuitive parenting. *Journal of Robotics and Mechatronics* 19(3), 2007, 315-323.
- [18] M. Otake, K. Arai, M. Kato, T. Maeda, Y. Ikemoto, K. Kawabata, T. Takagi, H. Asama, Experimental Analysis of the Attribution of Own Actions to Intention of Self or Others by the Multiple Forward Models. *Journal of Robotics and Mechatronics*, 19(4), 2007, 482-488.
- [19] Y. D. Sato, Y. Kuriya, Efficient and Effective Gabor Feature Representation for Face Detection. *International Journal of Computer and Information Engineering* 6, 2012, 9-12.
- [20] P. Viola, M. J. Jones, Robust Real-Time Face Detection. *Int. J. Computer Vision* 57(2), 137-154, 2004.

Crypto-Arithmetic Problem using Parallel Genetic Algorithm (PGA)

Mr. Shedge Kishor Namdeo, Mr. Sravan Kumar G., Miss. Pooja Kuyate

¹(M-Tech Scholar, CSE, RGPV/ LNCT/Indore, Indore, MP.)

²(Department of Computer, RGPV/ LNCT/ Indore, Indore MP.)

³(Department of Computer, University of Pune/ SVIT/ PREC, Loni., Nashik, Maharashtra.)

Abstract: Cryptarithmic puzzles are quite old and their inventor is not known. An example in *The American Agriculturist* of 1864 disproves the popular notion that it was invented by Sam Loyd. The name cryptarithmic was coined by puzzlist Minos (pseudonym of Maurice Vatriquant) in the May 1931 issue of *Sphinx*, a Belgian magazine of recreational mathematics. In the 1955, J. A. H. Hunter introduced the word "alphabetic" to designate cryptarithms, such as Dudeney's, whose letters from meaningful words or phrases. Solving a cryptarithm by hand usually involves a mix of deductions and exhaustive tests of possibilities. Cryptarithmic is a puzzle consisting of an arithmetic problem in which the digits have been replaced by letters of the alphabet. The goal is to decipher the letters (i.e. Map them back onto the digits) using the constraints provided by arithmetic and the additional constraint that no two letters can have the same numerical value.

Keywords: Genetic algorithms, Parallel Processing, Scheduling, Cryptarithmic, Parallel Genetic Algorithms.

I. INTRODUCTION

Cryptarithm is a genre of mathematical puzzle in which the digits are replaced by letters of the alphabet or other symbols. Cryptarithmic is the science and art of creating and solving cryptarithms. The world's best known Cryptarithmic puzzle is undoubtedly the puzzle shown in Figure 1. This was first introduced by H.E. Dudeney and was first published in the July 1924 issue of *Strand Magazine* associated with the story of a Kidnapper's ransom demand [10].

Modernization, by introducing computers and the Internet, is making quite an impact on Cryptarithmic and it has already become a standard AI problem because it characterizes a number of important problems in computer science arena. A rule based searching technique can provide the solution in minimum time.

$$\begin{array}{r} \text{S E N D} \\ + \text{M O R E} \\ \hline \text{M O N E Y} \end{array}$$

Figure 1: Cryptarithmic Puzzle

Cryptarithmic is a class of constraint satisfaction problems which includes making mathematical relations between meaningful words using simple arithmetic operators like 'plus' in a way that the result is conceptually true, and assigning digits to the letters of these words and generating numbers in order to make correct arithmetic operations as well[14].

GENETIC ALGORITHM:

Genetic algorithms were formally introduced in the United States in the 1970s by John Holland at University of Michigan. The continuing price/performance improvement of computational systems has made them attractive for some types of optimization. In particular, genetic algorithms work very well on mixed, combinatorial problems. They are less susceptible to getting 'stuck' at local optima than gradient search methods. But they tend to be computationally expensive. To use a genetic algorithm, you must represent a solution to your problem as a Chromosome. The genetic algorithm then creates a population of solutions and applies genetic operators such as mutation and crossover to evolve the solutions in order to find the best one(s)[4].

MOTIVATION

Cryptarithmic is a class of constraint satisfaction problems which includes making mathematical relations between meaningful words using simple arithmetic operators like 'plus' in a way that the result is conceptually true, and assigning digits to the letters of these words and generating numbers in order to make correct arithmetic operations as well. A simple way to solve such problems is by depth first search (DFS) algorithm which has a big search space even for quite small problems. I am proposing a solution to this problem with genetic algorithm and then optimized it by using parallelism. I also showed that the algorithm reaches a solution faster and in a smaller number of iterations than similar algorithms.

OBJECTIVES

In the beginning, there are randomly generated individuals. All those individuals create a population. The population in certain time is called a generation. According to their qualities they are chosen by operators for creation of a new generation. The quality of the population grows or decreases and give limits to some constant. Every individual is represented by its chromosome. Mostly chromosomes represented as a binary string. Sometimes there are more strings which are not necessarily of a binary type. The chromosome representation could be evaluated by a fitness function. The fitness equals to the quality of an individual and is an important pick factor for a selection process. The average fitness of a population changes gradually during the run. Operating on the population, several operators are defined. After choosing randomly a pair of individuals, crossover executes an exchange of the substring within the pair with some probability. There are many types of crossovers defined, but a description is beyond the scope of this report. Mutation is an operator for a slight change of one individual/several individual in the population. It is random, so it is against staying in the local minimum. Low mutation parameter means low probability of mutation. Selection identifies the fittest individuals. The higher the fitness, the bigger the probability to

become a parent in the next generation. The computation time for serial GA execution becomes high for time consuming fitness functions such as those including finite element analysis (FEA) at each objective function call. A better alternative is to take advantage of the intrinsically parallel nature of GAs and to perform the generation of new populations in parallel, on different processors.

II. LITERATURE REVIEW/ SURVEY

Cryptarithmic is a puzzle consisting of an arithmetic problem in which the digits have been replaced by letters of the alphabet. The goal is to decipher the letters (i.e. Map them back onto the digits) using the constraints provided by arithmetic and the additional constraint that no two letters can have the same numerical value.

Cryptarithmic is a class of constraint satisfaction problems which includes making mathematical relations between meaningful words using simple arithmetic operators like 'plus' in a way that the result is conceptually true, and assigning digits to the letters of these words and generating numbers in order to make correct arithmetic operations as well[1].

CONSTRAINT SATISFACTION PROBLEM

Cryptarithmic is a suitable example of Constraint Satisfaction Problem. Instead of providing description, a Cryptarithmic problem can be better described by some constraints [12].

Constraints of the Cryptarithmic problem are as follows:

- The arithmetic operations are in decimal; therefore, there must be maximum ten different letters in overall strings which are being used.
- All of the same letters should be bound to a unique digit and no two different letters could be bounded to the same digit.
- As the words will represent numbers, the first letter of them could not be assigned to zero.
- The resulting numbers should satisfy the problem, meaning that the result of the two first numbers (operands) under the specified arithmetic operation (plus operator) should be the third number.

Consider that, the base of the numbers is 10. Then there must be at most 10 unique symbols or letters in the problem. Otherwise, it would not be possible to assign a unique digit to each unique letter or symbol in the problem. To be semantically meaningful, a number must not begin with a zero. So, the letters at the beginning of each number should not correspond to zero.

WHY GENETIC ALGORITHMS?

It is better than conventional AI in that it is more robust. Unlike older AI systems, they do not break easily even if the inputs changed slightly, or in the presence of reasonable noise. Also, in searching a large state-space, multi-modal state-space, or n-dimensional surface, a genetic algorithm may offer significant benefits over more typical search of optimization techniques (linear programming, heuristic, depth-first, breath-first.)[15]. A genetic algorithm (GA) is a search technique used in computing to find exact or approximate solutions to optimization and search problems. Genetic algorithms are a type of iterative mathematical modeling

technique used to find the optimal combinatorial state given a set of parameters of interest.

Genetic algorithms (GAs) are powerful search techniques that are used successfully to solve problems in many different disciplines. Parallel GAs are particularly easy to implement and promise substantial gains in performance and as such there has been extensive research in this field. Genetic algorithms are based on natural selection discovered by Charles Darwin. They employ natural selection of fittest individuals as optimization problem solver. Optimization is performed through natural exchange of genetic material between parents. Offspring's are formed from parent genes. Fitness of offspring's is evaluated. The fittest individuals are allowed to breed only. Offspring's are created during crossover and mutation. The crossover is an operation when new Chromosomes offspring's are produced by fusing parts of other chromosomes parents. The mutation is random replacement of chromosome bits. Thus, offspring's form a new generation which replaces the old one.

The success of optimization strongly depends on the chosen chromosome encoding scheme, crossover and mutation strategies as well as fitness function. For each problem, careful analysis must be done and correct approach chosen. As it was shown, one Chromosome can contain a whole image or only a small part of it, a whole parameter range or only the most descriptive ones. Crossover can be performed in various manners, for example by exchanging information at one brake point or at several one.

III. PARALLEL GENETIC ALGORITHMS

If we mimic natural evolution we would not operate on a single population in which a given individual has the potential to mate with any other partner in the entire population. Instead, species would tend to reproduce within subgroups or within neighborhoods. A large population distributed among a number of semi-isolated breeding groups is known as polytypic. A PGA introduces the concept of interconnected demes. The local selection and reproduction rules allow the species to evolve locally, and diversity is enhanced by migrations of strings among demes [13].

In a genetic algorithm, a population of strings (called chromosomes or the genotype of the genome), which encode candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem, evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached. Genetic algorithms find application in bioinformatics, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics and other fields[4].

A TYPICAL GENETIC ALGORITHM REQUIRES:

A genetic representation of the solution domain,

A fitness function to evaluate the solution domain.

A standard representation of the solution is as an array of bits. Arrays of other types and structures can be used in essentially the same way. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size, which facilitates simple crossover operations. Variable length representations may also be used, but crossover implementation is more complex in this case. Tree like representations are explored in genetic programming and graph form representations are explored in evolutionary programming. Evolutionary programming (EP) involves populations of solutions with primarily mutation and selection and arbitrary representations. They use self adaptation to adjust parameters, and can include other variation operations such as combining information from multiple parents.

The fitness function is defined over the genetic representation and measures the quality of the represented solution. The fitness function is always problem dependent. For instance, in the knapsack problem one wants to maximize the total value of objects that can be put in a knapsack of some fixed capacity. A representation of a solution might be an array of bits, where each bit represents a different object, and the value of the bit (0 or 1) represents whether or not the object is in the knapsack. Not every such representation is valid, as the size of objects may exceed the capacity of the knapsack. The fitness of the solution is the sum of values of all objects in the knapsack if the representation is valid or 0 otherwise. In some problems, it is hard or even impossible to define the fitness expression; in these cases, interactive genetic algorithms are used.

Once the genetic representation and the fitness functions are defined, GA proceeds to initialize a population of solutions randomly, and then improve it through repetitive application of mutation, crossover, inversion and selection operators.

Initialization

Initially many individual solutions are randomly generated to form an initial population. The population size depends on the nature of the problem, but typically contains several hundreds or thousands of possible solutions. Traditionally, the population is generated randomly, covering the entire range of possible solutions (the search space). Occasionally, the solutions may be "seeded" in areas where optimal solutions are likely to be found.

Selection (genetic algorithm)

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a fitness based process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions. Other methods rate only a random sample of the population, as this process may be very time consuming.

Most functions are stochastic and designed so that a small proportion of less fit solutions are selected. This helps keep the diversity of the population large, preventing premature convergence on poor solutions. Popular and well

studied selection methods include roulette wheel selection and tournament selection.

Reproduction

Crossover (genetic algorithm) and Mutation (genetic algorithm)

The next step is to generate a second generation population of solutions from those selected through genetic operators: crossover (also called recombination), and/or mutation. For each new solution to be produced, a pair of "parent" solutions is selected for breeding from the pool selected previously. By producing a "child" solution using the above methods of crossover and mutation, a new solution is created which typically shares many of the characteristics of its "parents". New parents are selected for each new child, and the process continues until a new population of solutions of appropriate size is generated. Although reproduction methods that are based on the use of two parents are more "biology inspired", some research suggests more than two "parents" are better to be used to reproduce a good quality chromosome.

These processes ultimately result in the next generation population of chromosomes that is different from the initial generation. Generally the average fitness will have increased by this procedure for the population, since only the best organisms from the first generation are selected for breeding, along with a small proportion of less fit solutions, for reasons already mentioned above. Although, Crossover and Mutation are known as the main genetic operators, it is possible to use other operators such as regrouping or migration in genetic algorithms.

- Simple generational genetic algorithm pseudo code:
- Choose the initial population of individuals.
- Evaluate the fitness of each individual in that population.
- Repeat on this generation until termination: (time limit, sufficient fitness achieved, etc.)
- Select the best-fit individuals for reproduction .
- Breed new individuals through crossover and mutation operations to give birth to offspring.
- Evaluate the individual fitness of new individuals.
- Replace least-fit population with new individuals.

Genetic algorithms with adaptive parameters (adaptive genetic algorithms, AGAs) is another significant and promising variant of genetic algorithms. The probabilities of crossover (pc) and mutation (pm) greatly determine the degree of solution accuracy and the convergence speed that genetic algorithms can obtain. Instead of using fixed values of pc and pm, AGAs utilize the population information in each generation and adaptively adjust the pc and pm in order to maintain the population diversity as well as to sustain the convergence capacity. In AGA (adaptive genetic algorithm), the adjustment of pc and pm depends on the fitness values of the solutions. In CAGA (clustering based adaptive genetic algorithm), through the use of clustering analysis to judge the optimization states of the population, the adjustment of pc and pm depends on these optimization states. It can be quite effective to combine GA with other optimization methods. GA tends to be quite good at finding generally good global solutions, but quite inefficient at finding the last few mutations to find the absolute optimum. Other techniques (such as simple hill climbing) are quite efficient at finding absolute optimum in a limited region. Alternating GA and hill climbing can

improve the efficiency of GA while overcoming the lack of robustness of hill climbing [6].

Genetic operators as independent parts of GA

The parallel steady-state genetic algorithm with tournament bad individual selection was implemented. In this implementation [6] the genetic algorithm consists of two threads: one performs tournament selection and crossover and the other mutation. The major problem of that simple parallel implementation is that it has no control over mutation probability. The consequence is a very bad algorithm behavior. The results are slightly better than random search, but also useless. If the threads are left to parallel execution without any control, one of two threads can waste some time on waiting for processor time.

An Evolutionary Algorithm will search for solutions in shortest time but the performance will also reflect the toughness of the problem. A parallel genetic algorithm has been developed to dynamically schedule heterogeneous tasks to heterogeneous processors in a distributed environment. The proposed algorithm uses multiple processors with centralized control for scheduling. Tasks are taken as batches and are scheduled to minimize the execution time and balance the loads of the processors.

IV. SUMMARY & DISCUSSION

In this project we try to analyze an efficient Parallel genetic algorithm to solve Cryptarithmic Problems. Additionally, it illustrates how to plug in techniques of Evolutionary Approach into Constraint Satisfaction Problem. This sort of design can provide efficient solution to a wide range of Constraint Satisfaction Problem or other generic searching problems that could be characterized as a Constraint Satisfaction Problem as well. This parallel model has been tested in order to determine the best method for comparing, science it uses two platform-independent parameters; the number iteration and java programming language. So, further research should go on to optimize the main proposed parallel ideas in the near future. This project concentrated on solving Cryptarithmic problems in an efficient way. The use of parallel genetic algorithm showed that we can even find the result of large instances of this problem within an acceptable time.

Discussion is related a simple Cryptarithmic problem solution in stepwise mode - Cryptarithmic is a CSP problem in which letters are substituted by digits such that each letter represents a unique digit, and the actual problem is to find a proper sequence of digits assigned to different letters satisfying the conditions of the arithmetic operation. What is a Cryptarithmic problem? It is a mathematical puzzle in which each letter represents a digit (for example, if $X=3$, then $XX=33$). The object is to find the value of each letter. No two letters represent the same digit (If $X=3$, Y cannot be 3). And the first letter cannot be 0 (Given the value ZW , Z cannot be 0). They can be quite challenging, often involving many steps. Here's an example, illustrating how to solve them:

$$\begin{array}{r} \text{SEND} \\ + \text{MORE} \\ \hline \text{MONEY} \end{array}$$

M must be 1. This is an addition problem; the sum of two four digit numbers can't be more than 10,000, and M can't be 0 according to the rules since it's the first letter. So now we have:

$$\begin{array}{r} \text{SEND} \\ + \text{1ORE} \\ \hline \text{1ONEY} \end{array}$$

Now in the column S1O, $S+1 \geq 10$. S must be 8 (if there is a 1 carried over from the column E0N) or 9. O must be 0 (if $S=8$ and there is a 1 carried or $S=9$ and there is no 1 carried) or 1 (if $S=9$ and there is a 1 carried). But 1 is already taken, so O must be 0.

$$\begin{array}{r} \text{SEND} \\ + \text{10RE} \\ \hline \text{1ONEY} \end{array}$$

There can't be a carry from the column E0N, because any digit plus $0 < 10$, unless there is a carry from the column NRE and $E=9$; but this cannot be the case, because then N would be 0, and 0 is already taken. So $E < 9$ and there is no carry from this column. Therefore, $S=9$, because $9+1=10$.

In the column E0N, E cannot be equal to N, so there must be a carry from the column NRE; $E+1=N$. We now look at the column NRE; we know that $E+1=N$. Since we know that there is a carry from this column, $N+R=1E$ (if there is no carry from the column DEY) or $N+R+1=1E$ (if there is a carry from the column DEY). Let's try out both cases. No carry: $N+R=10+(N-1)=N+9$,

$R=9$; 9 is already taken, so this won't work.

Carry: $N+R+1=N+9$; $R=8$. This must be the solution for R.

$$\begin{array}{r} \text{9END} \\ + \text{108E} \\ \hline \text{10NEY} \end{array}$$

The digits we have left are 7, 6, 5, 4, 3, and 2. We know there must be a carry from the column DEY, so $D+E > 10$. $N=E+1$, so E can't be 7 because then N would be 8 which is already taken. D is at most 7, so E cannot be 2 because then $D+E < 10$, and E cannot be 3 because then $D+E=10$ and $Y=0$, but 0 is taken already. Likewise, E cannot be 4 because if $D > 6$, $D+E < 10$, and if $D=6$ or $D=7$, then $Y=0$ or $Y=1$, which are both taken. So E is 5 or 6. If $E=6$, then $D=7$ and $Y=3$, so this part works. But look at the column N8E. Remember, there is a carry from the column D5Y. $N+8+1=16$ (because we know there is a carry for this column). But then $N=7$, and 7 is taken by D. Therefore, $E=5$.

$$\begin{array}{r} \text{95ND} \\ + \text{1085} \\ \hline \text{10N5Y} \end{array}$$

Now that we've gotten this important digit, it gets much simpler from here. $N+8+1=15$, $N=6$.

956D
+ 1085

1065Y

The digits left are 7, 4, 3, and 2. We know there is a carry from the column D5Y, so the only pair that fits is $D=7$ and $Y=2$

9567
+ 1085

10652

The problem is solved! These are quite tricky and require some thinking, but are lots of fun. Now we'll take turns posting problems. When a problem is solved.

V. CONCLUSION

In Cryptarithmic puzzle, the arithmetic operations are simple and of base 10, but are ciphered into letters. The task is to decipher them. Here we concentrated on solving Cryptarithmic problems in an efficient way. Parallel implementations of Genetic Algorithms are very performable to solve large scale problems. The use of parallel genetic algorithm showed that we can even find the result of large instances of this problem within an acceptable time. The proposed algorithm uses multiple processors with centralized control for scheduling. Tasks are taken as batches and are scheduled to minimize the execution time and balance the load among of the processors. A scheduling algorithm has been developed to schedule heterogeneous tasks onto heterogeneous processors on a distributed environment.

Genetic Algorithms are powerful but usually suffer from longer scheduling time which is reduced in our algorithm due to the parallelization of the fitness evaluation. The proposed algorithm uses a straightforward encoding scheme and generates a randomized initial population. The fitness function uses the maxspan, balance of load among the processors and communication costs while evaluating the schedules. By parallelization I got a better program structure and a significant decrease in computational time on a multiprocessor system. As per the implementation, testing, result analysis I conclude that PGA and DFS implementation is 80 to 90 % successful.

REFERENCES

[1]. Abu Sayef Md. Ishaque, Md. Bahlul Haider, Muhammad Al Mahmud Wasid, Shah Mohammed Alaul, Md. Kamrul Hassan, Tanveer Ahsan, Md. Shamsul Alam: "An Evolutionary Algorithm to Solve Cryptarithmic Problem", International Conference on Computational Intelligence 2004: 494-496.

[2]. AJ Page, TJ Naughton "Dynamic Task Scheduling using Genetic Algorithms for Heterogeneous Distributed Computing", - Proceedings of the 19th

IEEE International Parallel and Distributed Processing Symposium (IPDPS'05)

[3]. A. Y. Zomaya, M. Clements, and S. Olariu. "A framework for Reinforcement-Based Scheduling in Parallel Processor Systems", IEEE Transactions on Parallel and Distributed Systems, 9(3):249-260, March 1998.

[4]. Alippi, C., Filho, J.L.R., Treleaven, P.C. (1994), "Genetic-Algorithm Programming Environments", IEEE Trans. Computer, June 1994.

[5]. Bonnie Averbach and Orin Chein, Problem Solving Through Recreational Mathematics, Courier Dover Publications, 2000, pp. 156.

[6]. Budin, L., Golub, M., Jakobovic, D., Parallel Adaptive Genetic Algorithm, International ICSC/IFAC Symposium on Neural Computation NC'98, Vienna, 1998, pp. 157-163.

[7]. Cantú-Paz, E., "Designing Efficient Master Slave Parallel Genetic Algorithms, Genetic Programming": Proceedings of the Third Annual Conference. (pp. 455). San Francisco, CA, 1998.

[8]. Cantú-Paz E., A Summary of Research on Parallel Genetic Algorithms, 1995., available from: www.dai.ed.ac.uk/groups/evalg/Local_Copies_of_Paper_s/Cantu-Paz.A_Summary_of_Research_on_Parallel_Genetic_Algorithms.ps.gz

[9]. Cantú-Paz, E., A Survey of Parallel Genetic Algorithms, Calculateurs Paralleles, Vol. 10, No. 2. Paris: Hermes, 1998., available via ftp from: <ftp://ftp-illigal.ge.uiuc.edu/pub/papers/Publications/cantupaz/survey.ps.Z>.

[10]. Cantú-Paz, E., Designing Efficient Masterslave Parallel Genetic Algorithms, Genetic Programming: Proceedings of the Third Annual Conference. (pp. 455). San Francisco, CA, 1998.

[11]. Cantú-Paz, E., Goldberg, D.E., Parallel Genetic Algorithms with Distributed Panmictic Populations, 1999. available from: <http://wwwilligal.ge.uiuc.edu/cgi-bin/orderform/orderform.cgi>.

[12]. Cantú-Paz, E., Migration Policies, Selection Pressure, and Parallel Evolutionary Algorithms, 1999., available via ftp from: <ftp://ftpilligal.ge.uiuc.edu/pub/papers/IlliGALs/99015.ps.Z>.

[13]. David Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, MA, 1989.

[14]. E. Hou, N. Ansari, and H. Ren. "A Genetic Algorithm for Multiprocessor Scheduling". IEEE Transactions on Parallel and Distributed Systems, 5(2):113-120, February 1994.

[15]. E Cantu-Paz ,A survey of parallel genetic algorithms, Calculateurs Paralleles, Reseaux et Systems Repartis, 1998

[16]. Goodman, E.D., Averill, R.C., Punch, W.F., Eby, D.J., Parallel Genetic Algorithms in the Optimization of Composite Structures, Second World Conference on SoftComputing (WSC2), June, 1997., available from: <http://garage.cps.msu.edu/papers/GARAGe97-05-02.ps>

[17]. H. E. Dudeney, in Strand Magazine vol. 68 (July 1924), pp. 97 and 214.

- [18]. M. Maheswaran, S. Ali, H. J. Siegel, D. Hensgen, and R. F. Freund. "Dynamic Mapping of a Class of Independent Tasks onto Heterogeneous Computing Systems." *Journal of Parallel and Distributed Computing*, 59(2):107– 131, November 1999.
- [19]. Michalewicz, Z. (1992), *Genetic Algorithms + Data Structures = Evolutionary Programs*, Springer-Verlag, Berlin.
- [20]. R. Gholamali , D. Gholamhossein : "An Efficient Parallel Algorithm for Solving Cryptarithmic Problems: PGA", Third UKSim European Symposium on Computer Modeling and Simulation 2009.
- [21]. R.Nedunchelian, K.Koushik, N.Meiyappan, V.Raghu, "Dynamic Task Scheduling Using Parallel Genetic Algorithms for Heterogeneous Distributed Computing", *Proceedings of the 2006 International Conference on Grid, Las Vegas, Nevada, USA, 2006*.
- [22]. R. Abbasian, M. Mazloom : "Solving Cryptarithmic Problems Using Parallel Genetic Algorithm" 2009 Second International Conference on Computer Engineering 978-0-7695-3925-6/09 2009 IEEE
- [23]. Stuart Russell, *Artificial Intelligence A Modern Approach*, Pearson Education, Inc, ISBN 0-13-790395-2, 2003.
- [24]. SunSoft, (1994), *Solaris 2.4: Multithreaded Programming Guide*, Sun Microsystems, Mountain View, California
- [25]. Manisha Neaghare, "Solving Verbal Arithmetic Problem by Efficient Evolutionary Algorithm" *International Conference on Sunrise Technology iCOST2011-Computer Engg. Organised by SSVPS Engineering College, Dhule*.
- [26]. Manisha Neaghare "Comparison of Parallel Genetic Algorithm with Depth First Search Algorithm for Solving Verbal Arithmetic Problems" *International Conference on Emerging Trends in Technology ICWET2011 Organised by Thakur Engineering College, Mumbai published on Association for Computing Machinery ACM SIGART USA and International Journal of Computer Application*.
- [27]. Shedge Kishor N., Ashwini Verma, Gade Shyam A. "Solving Verbal Crypto-Arithmetic Problem by Parallel Genetic Algorithm (PGA)" *International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 2, Issue 4, August 2012*

Mass Transfer with Equilateral Triangular Plate as Turbulence Promoter in Circular Pipe

D. S. Seetharama Raju^a, M. Gangadhar^b, V. Nageswara Rao^a,
N. Chitti Babu^{a*}, P. J. Rao^a

^aDepartment of Chemical Engineering, College of Engineering(A), Andhra University, Andhra Pradesh, INDIA

^bDepartment of Chemical Engineering, G.M.R. Institute of Technology, Rajam, Andhra Pradesh, INDIA

ABSTRACT: Mass transfer data were obtained in circular conduit using an electrochemical technique with a potassium ferri-ferro cyanide couple. In circular conduit coaxially placed entry region equilateral triangular plate was used as turbulence promoter. The study comprised of evaluation of mass transfer rates at the outer wall of the electrochemical cell. Mass transfer coefficients were evaluated from the measured limiting currents. The study covered a wide range of geometric parameters such as side of the plate (S_p), thickness of the plate (T_p) and distance of the plate from the entrance of the test section (h). The results revealed that the mass transfer coefficient increased with increase in velocity, side of the plate (S_p), thickness of the plate (T_p) and decreased with increase in distance of the plate from the entrance of the test section (h). Within the range of variables covered, the augmentation achieved in mass transfer coefficients were up to 5.2 fold over the tube flow in absence of promoter. The entire mass transfer data were correlated with $g(h^+)$ and roughness Reynolds number. The following correlation was reported out of the study.

$$g(h^+) = 1.0308 (Re^+)^{0.0525} (\phi_1)^{0.2037} (\phi_2)^{0.16571} (\phi_3)^{-0.0527} (Sc)^{0.3227}$$

Where $\phi_1 = S_p/d$, $\phi_2 = T_p/d$, $\phi_3 = h/d$ are dimensionless groups, d is diameter of test section.

Keywords: Mass transfer, equilateral triangular plate, turbulence promoter

I. INTRODUCTION

Several researchers carried out extensive work to identify and establish different type of plate assemblies for enhancing heat, mass and momentum transfers in the reactors. In the operation of electrochemical reactor cells in the fields of electro-winning, refining, electro organic synthesis, etc., the significant importance of the cell plate characteristics, like its roughness, geometry, etc., has been well recognized. Lin et al [1] studied mass transfer rates with Ferri-Ferrocyanide systems in diffusion controlled reactions and Bergles [2] conducted an extensive review to evaluate an appropriate technique for improving heat and mass transfer. Several studies were also conducted by inserting string of spheres [3], string of discs [4], spiral coils [5], coaxially placed cones [6], axial rods with twisted tapes mounting [7], discs [8], orifices [9], across the flows in the conduits for increasing mass and momentum transfer. Coaxially placed tape-disc assembly was reported as a good turbulence promoter, thereby increasing mass transfer [10]. However, effects of coaxially placed equilateral triangular plate in circular conduits on mass transfer rates in case of forced convection flow of electrolyte have not been studied earlier and accordingly the present study was undertaken by the author to evaluate the benefits of turbulence on mass transfer by utilizing equilateral triangular plate placed coaxially in circular conduits and mass transfer rates at the conduit wall, by applying the limiting current technique. Pressure drop measurements were also carried out simultaneously to obtain data for computing power losses in the system.

Table 1 shows the parameters covered in the present study.

II. EXPERIMENTAL

Schematic diagram of experimental set up is shown in figure 1. It is similar in layout to that used in earlier studies [9, 10]. It essentially consisted of a storage tank (TS), centrifugal pump (P), rotameter (R), entrance calming section (E_1), test section (T) and exit calming section (E_2). The storage tank is cylindrical copper vessel of 100 liter capacity with a drain pipe and a gate valve (V_1) for periodical cleaning. A copper coil (H) with perforations is provided to bubble nitrogen through the electrolyte. The tank is connected to the pump with a 0.025m diameter copper pipe on the suction line of the centrifugal pump. The suction line is also provided with a gate valve (V_2). The discharge line from the pump splits into two. One served as a bypass line and controlled by valve (V_3). The other connects the pump to the entrance calming section (E_1) through rotameter. The rotameter is connected to a valve (V_4) for adjusting the flow at the desired value. The rotameter has a range of 0 to $166 \times 10^{-5} \text{m}^3/\text{s}$. The entrance calming section consisted of 0.05 m ID circular copper pipe with a flange and is closed at the bottom with a gland nut (G). The up-stream side of the entrance calming section is filled with capillary tubes to damp the flow fluctuations and to facilitate steady flow of the electrolyte through the test section. It is made of a graduated Perspex tube of 0.36m length with point electrodes fixed flush with the inner surface of the tube. The point electrodes are made out of a copper rod and machined to the size. They are fixed flush with the inner surface of the test section at equal spacing of 0.01m. Exit calming section is also of the same diameter copper tube of 0.5 m long, and it is provided with a flange on the upstream side for assembling the test section. It has gland nuts (G) at the top and bottom ends to hold the central tube. Two thermo wells (t_1, t_2) were provided, one at upstream side of the entrance calming section and the other at the down stream side of exit calming section for measurement of temperature of the electrolyte. Equilateral triangular plate serving as turbulence promoter is made of Nylon of various sizes with a provision to fix it rigidly within the test section. The plate is

placed concentrically in the test section. The promoters used are shown in photograph in figure 2. The details of equilateral triangular plate promoter are shown in figure 3. The limiting current measuring equipment consisted of multimeter of Motwane make which has 0.01mA accuracy and vacuum tube voltmeter is used for potential measurements. The other equipments used in circuit are rheostat, key, commutator, selector switch, and a lead acid battery as the power source. The commutator facilitated the measurement of limiting currents for oxidation and reduction process under identical operating conditions by the change of polarity while the selector switch facilitated the measurements of limiting currents at any desired electrode. The circuit diagram used for the measurement of limiting currents is shown in the figure 4.

Data on limiting currents for the case of reduction of ferricyanide ion is obtained for fluid flow in circular conduits in the presence of triangular plate as insert promoter. The following electrode reaction is involved.

Cathodic reduction of ferricyanide ion:



Initially blank runs are conducted with indifferent electrolyte (sodium hydroxide solution) alone to ensure that the limiting currents obtained in the subsequent runs are due to diffusion of reacting ions (Ferri cyanide ion) only. The electrolyte was pumped at a desired flow rate (through the test section) by operating the control and by-pass valves. After steady state is attained, potentials are applied across the test electrode and wall electrode in small increments of potentials (100mV) and the corresponding currents were measured for each increment. In view of the large area of the counter electrode in relation to the test electrode nearly constant potential is maintained at the test electrode. Since the potential values are not of criteria in the present study, the limiting currents were determined from the measurements of applied potentials and currents as has been done in several earlier works [9, 10]. The attainment of limiting current is indicated by the constancy of current with a large increase in the potential. Mass transfer coefficients are computed from the measured limiting currents by the following equation:

$$k_L = i_L / nFAC_0 \quad \dots\dots(2)$$

Pressure drop measurements are also taken simultaneously using U-tube manometer with Carbon tetrachloride as manometric liquid.

III. RESULTS AND DISCUSSION

Introduction of co-axially placed equilateral triangular plate in a circular conduit alters the flow pattern in cell by generating eddies and wakes. These eddies generally consume large amount of energy because of their circulatory nature. The eddies will in turn influence shear forces near the wall region which in turn reduce the thickness of the concentration boundary layer, there by augmenting mass transfer rates. At higher flow rates, velocity through constriction between plate and wall of the test section dominates, while the energy utilized for the wakes and eddies becomes marginal and thereby, most of the energy is utilized. Table 2 indicates the exponent on velocity of the present study together with the other works. The exponent on velocity is comparable to that in the studies on mass transfer with different turbulence generating systems.

IV. Effect of geometric parameters

Effect of side of equilateral triangular plate (S_p):

Side of triangular plate has strong influence on mass transfer coefficient k_L . k_L versus velocity of electrolyte (V) is drawn for different sides of plate and is shown in figure 5. Various sides of triangular plate used in the present study are $S_p=0.025\text{m}$, 0.030m , 0.035m , 0.040m . Mass transfer coefficient increases with increase in side of triangular plate. The augmentation in mass transfer coefficient is 1.04 times over the smooth tube [1] for side of plate 0.025m at a velocity of 0.3936m/s while the augmentation is 1.64 times over the smooth tube [1] for the side 0.040m at the same velocity of 0.3936m/s .

Effect of plate thickness (T_p):

In figure 6, mass transfer coefficient k_L is drawn against velocity to study the effect of plate thickness on mass transfer by keeping all the other parameters constant. The thickness of the plate used in the present study are $T_p=0.005\text{m}$, 0.020m , 0.040m , 0.060m . Mass transfer coefficient increases from 1.52 times to 5.23 times as the thickness of the plate increases from 0.005m to 0.060m at the velocity of 0.3936 m/s .

Effect of location of plate inside the test section (h):

As the distance of the plate from the entrance of the test section varied, extent of turbulence also varied because of change in circulating pattern and it is extended to both plate region and down stream region to plate region. Mass transfer coefficient (k_L) versus velocity of electrolyte for different distances of the plate(h) from the entrance of the test section are plotted and is shown in figure 7. Mass transfer coefficients are decreased from 1.8 to 1.23 over Lin et al [1] for the smooth tube, while the distance of the plate from the entrance of the test section increases from 0.14m to 0.30m

V. DEVELOPMENT OF CORRELATIONS

The data on mass transfer with equilateral triangular plate as turbulence promoter could well be calculated in the lines done in earlier studies [10]. Correlation of data using colburn J_D factor with Reynolds number yielded the following equation

$$J_D = 0.43641(\text{Re})^{-0.38926} \dots\dots\dots(3)$$

Average deviation = 58.55, Standard deviation = 87.56
 By incorporating dimensionless geometrical groups, the following correlation is yielded

$$J_D = 0.163 \times 10^{-3} (\text{Re})^{-0.3007} (\phi_1)^{0.9926} (\phi_2)^{0.4541} (\phi_3)^{-0.5566} (\text{Sc})^{1.3054} \dots\dots\dots(4)$$

Average deviation = 28.598, Standard deviation = 37.815

Where $\phi_1 = S_p/d$, $\phi_2 = T_p/d$, $\phi_3 = h/d$, Sc which are dimensionless groups.

The above equation show large deviation. So mass transfer data are calculated in the lines similar to earlier studies [10] by using roughness mass transfer function $g(h^+)$ in place of ' J_D ' and roughness Reynolds number Re^+ in place of Re . The side of plate S_p is chosen as effective geometric parameter.

$R(h^+)$ and Re^+ are defined as follows

$$R(h^+) = 2.5 \ln[2S_p/d] + \sqrt{(2/f)} + 3.75 \dots\dots\dots(5)$$

$$\text{Re}^+ = (S_p/d) \cdot \text{Re} \cdot \sqrt{(f/2)} \dots\dots\dots(6)$$

$$g(h^+) = S_t / S_{t_0} + R(h^+) \dots\dots\dots(7)$$

The following correlation is obtained without incorporating geometrical groups.

$$g(h^+) = 3.3294 (\text{Re}^+)^{0.1202} \dots\dots\dots(8)$$

Average deviation = 14.064, Standard deviation = 22.665

The following correlation is obtained by incorporating dimensionless geometrical groups and Schmidt number.

$$g(h^+) = 1.0308 (\text{Re}^+)^{0.0525} (\phi_1)^{0.2037} (\phi_2)^{0.16571} (\phi_3)^{-0.0527} (\text{Sc})^{0.3227} \dots\dots\dots(9)$$

Average deviation = 5.984, Standard deviation = 10.047

Correlation plot for equations (9) is presented in the figure 8.

VI. COMPARISON OF CORRELATIONS

For a selected set of geometric parameters correlation factor for mass transfer (Y_1) is plotted against Re^+ , for comparison data with other studies namely Rao [14]. Having comparable geometric parameters, are computed with present method is shown in figure 9. The data falls close to the present study indicating correlation presented in the present work is comparable.

$$Y_1 = g(h^+) / (S_p/d)^{0.2037} (T_p/d)^{0.1657} (h/d)^{-0.0527} (\text{Sc})^{0.3227} \dots\dots\dots(10)$$

VII. CONCLUSIONS

Mass transfer coefficients are increasing with increase in velocity. Mass transfer coefficients are increasing with increase in side of plate (S_p). Mass transfer coefficients are increasing with increases in thickness of the plate (T_p) and decreases as distance of the plate (h) from the entrance of the test section increases. In the present study, it is found that side of plate (S_p) = 0.040m, plate thickness (T_p) = 0.060m, distance of the plate from the entrance of the test section (h) = 0.14m, offered maximum augmentation. A maximum augmentation of 5.23 folds is observed over smooth tube flow without turbulence promoter. Correlations developed based on semi theoretical considerations. Wall similarity concept is applied for the present case.

Correlations developed for mass transfer:

$$g(h^+) = 1.0308 (\text{Re}^+)^{0.0525} (\phi_1)^{0.2037} (\phi_2)^{0.16571} (\phi_3)^{-0.0527} (\text{Sc})^{0.3227}$$

Dimensionless Groups:

- J_D = Mass Transfer Factor (k_L/V). $\text{Sc}^{2/3}$
- Re = Reynolds number = $dV\rho/\mu$
- Re^+ = Roughness Reynolds number = $(S_p / d) \cdot \text{Re} \cdot \sqrt{(f/2)}$
- $R(h^+)$ = Roughness momentum transfer function = $2.5 \ln(2 S_p / d) + \sqrt{(2/f)} + 3.75$
- St = Stanton number = k_L/V
- St_0 = Stanton number for conduit without internals
- Sc = Schmidt number $\mu/\rho D_L$
- Sh = Sherwood number $k_L d/D_L$
- u^+ = dimensionless velocity, u/u^*
- y^+ = dimensionless radial distance from the wall, $y u^*/\nu$
- $g(h^+)$ = $S_t / S_{t_0} + R(h^+)$

Nomenclature:

d	=	Diameter of test section, m
D_L	=	Diffusivity of reacting ion, m^2/sec
E	=	Energy consumed using triangular plate in the conduit, N/m^2
E_o	=	Energy consumed for empty conduit, N/m^2
f	=	Friction factor, $\Delta p d g_c / 2LV^2 \rho$
ΔP	=	Pressure difference, N/m^2
F	=	Faraday's constant = 96,500 coulombs/g-mol
g	=	Acceleration due to gravity, m/sec^2
g_c	=	Gravitational constant.
i_L	=	Limiting current, amp
k_L	=	Mass Transfer coefficient, m/s
k_o	=	Mass transfer coefficient of the empty conduit, m/s
L	=	Length of Test section, m
n	=	Number of electrons transferred
Q	=	Volumetric flow rate, m^3/s
S_t	=	Side of plate, m
T_t	=	Thickness of plate, m
h	=	Location of the plate from the entrance of the test section, m
u	=	Local velocity, m/s
u^*	=	Friction velocity = $\sqrt{(\tau_w g_c / \rho)}$, m/s
V	=	Average velocity, m/s
y	=	Radial distance from the wall, m
Y_1	=	$g(h^+) / (S_p/d)^{0.2037} (T_p/d)^{0.1657} (h/d)^{-0.0527} (Sc)^{0.3227}$

Greek letters:

ϵ	=	Eddy viscosity, m^2/s
ϵ_D	=	Eddy diffusivity, m^2/s
μ	=	Viscosity of fluid, $Kg/m \cdot sec$
ν	=	Kinematic viscosity, m^2/s
ρ_c	=	Density of manometer fluid, Kg/m^3
ρ	=	Density of fluid, Kg/m^3
τ_w	=	Shear stress, N/m^2

References

- [1]. Lin,C.S. Denton,E.B., Gaskill,H.S. and Putnam,G.L., "Diffusion controlled electrode reactions" *Ind.and Engg. Chemistry*, 43(9),1951, 2136-2143.
- [2]. Bergles, A.E., "Survey and evaluation of techniques to augment convective heat and mass transfer" *Progress in Heat and Mass transfer*, 1, 1969, 331-424.
- [3]. Sitaraman, T.S., "Augmentation of mass transfer with coaxially string of spheres as internals in tubes in fluidized beds", Ph.D Thesis., University of Madras, Madras, 1977.
- [4]. P.Venkateswarlu, T.Gopichand, G.J.V.J.Raju, "Increased mass transfer in a circular column in the presence of disc promoter", *J Energy Heat Mass Transfer*, 22, 2000, 195-203.
- [5]. Rajendra Prasad, P., "Studies on ionic mass transfer with coaxially placed spiral coils as turbulence promoter in homogenous flow and in fluidised beds", Ph.D Thesis, Andhra University, Visakhapatnam, India, 1993.
- [6]. S.Sarveswara Rao, G.J.V.J.Raju, "Studies on Ionic Mass Transfer with co-axially placed cones on a rod in homogeneous fluid and fluidized beds", *Paper presented at World Congress III of Chemical Engineers*, Tokyo, 1986, 481.
- [7]. V.Sujatha,C.Bhaskara Sarma,G.J.V.Jagannadha Raju, 'Studies on ionic mass transfer with coaxially placed helical tapes on a rod in forced convection flow",*Chem.Engg.Process*,36, 1997, 67-73.
- [8]. Evans, L.B. and Churchill, S.W., "Chem.Engg.Progr.Symp.Series 59, 36, *Ind. J.Technol.* 17, 1979, 245.
- [9]. Sujatha Kumari, B., "Studies on ionic mass transfer with coaxially placed orifice as turbulence promoter", M Tech Thesis, Andhra University, Visakhapatnam, India, 2003.
- [10]. Nageswara Rao, V., "Studies on ionic mass and momentum transfer with coaxially placed twisted tape –disc assembly as turbulence promoter in circular conduit", Ph.D Thesis Andhra University, Visakhapatnam, India, 2004.
- [11]. D.Changal Raju,C.Bhaskara Sharma,G.J.V.Raju, "Ionic mass transfer at the outer wall of concentric annuli in the presence of fluidizing solids-Effect of wires wound on central rod", *J.Energy Heat Masss Transfer*,14, 1992, 151-156.
- [12]. Klaczac,A., Trans. A.S.M.E., *J. Heat Transfer*,95, 1973, 557.

- [13]. V.Nageswara Rao and S. Teja Latha, "Mass transfer in circular tube with square-grooved serrated disc as turbulence Promoter", *Int. J. Chemical Engineering Research*, 2, 2010, 335-344.
- [14]. P J Rao "Studies on ionic mass and momentum transfer with coaxially placed square plate assembly as turbulence promoter in circular conduit", Ph.D Thesis Andhra University, Visakhapatnam, India, 2010.

Table: 1 Range of variables covered in the present study

Variable	Minimum	Maximum	Max/Min
Side of equilateral triangular plate, S_p , m	0.025	0.040	1.6
Thickness of equilateral triangular plate, T_p , m	0.005	0.060	12
Distance of the plate from entrance of the test section, h, m	0.14	0.30	2.14
Velocity V, m/s	0.0984	0.3936	4
Reynolds Number, Re	5264	21059	4
Schmidt Number, Sc	870.3	1040	1.19

Table 2:

Author	Promoter	System	Exponent on velocity	Range of Re
Klaczack [12]	Spiral coil	Mass transfer	0.520	1700-20000
Sujatha [7]	Tapes mounted on a rod	Mass transfer	0.490	1348-30605
Venkateswarlu [4]	String of discs	Mass transfer	0.498	3300-18650
Sitaraman [3]	String of spheres	Mass transfer	0.556	100-34000
Sarveswara Rao [6]	String of cones	Mass transfer	0.431	690-20200
Changal Raju [11]	Wires wound on a rod	Mass transfer	0.490	1500-20000
Nageswara Rao [10]	Tape-disc assembly	Mass transfer	0.485	1300-12000
Teja Latha et al [13]	Square grooved Serrated disc Square plate assembly	Mass transfer	0.205	1933-19337
P Jagannadha Rao[14]		Mass transfer	0.482	5856-15226
Present work	Equilateral triangular plate	Mass transfer	0.495	5264-21096

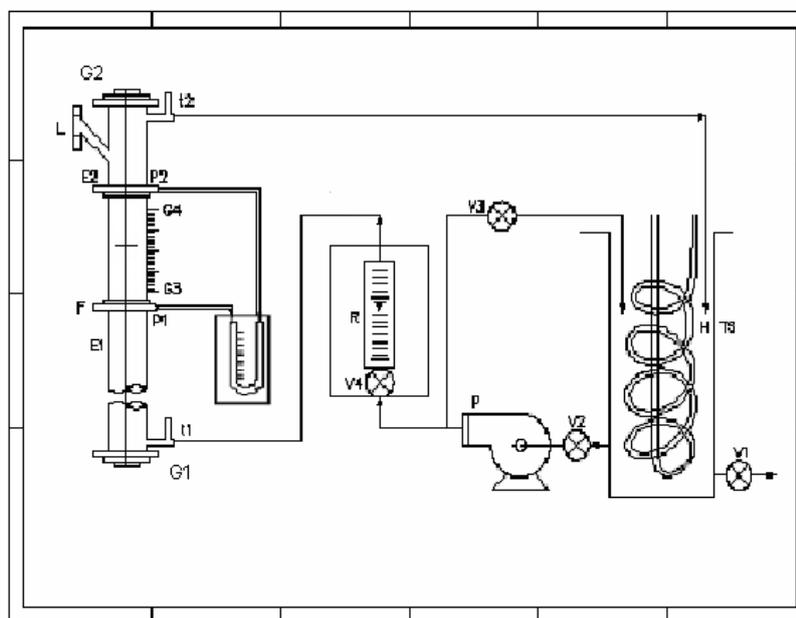


Figure 1: Schematic Diagram of Experimental Setup

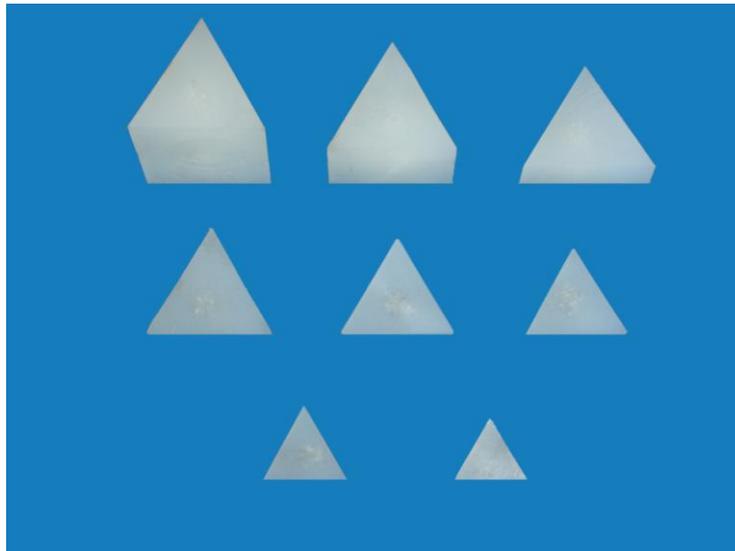
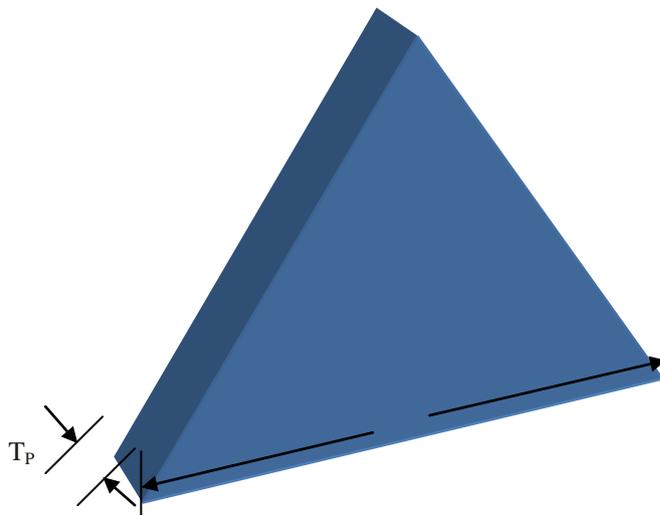


Figure 2: Turbulence Promoters



Note: S_p = Side of Triangular Plate, T_p = Thickness of Triangular Plate

Figure 3: Details of Promoter

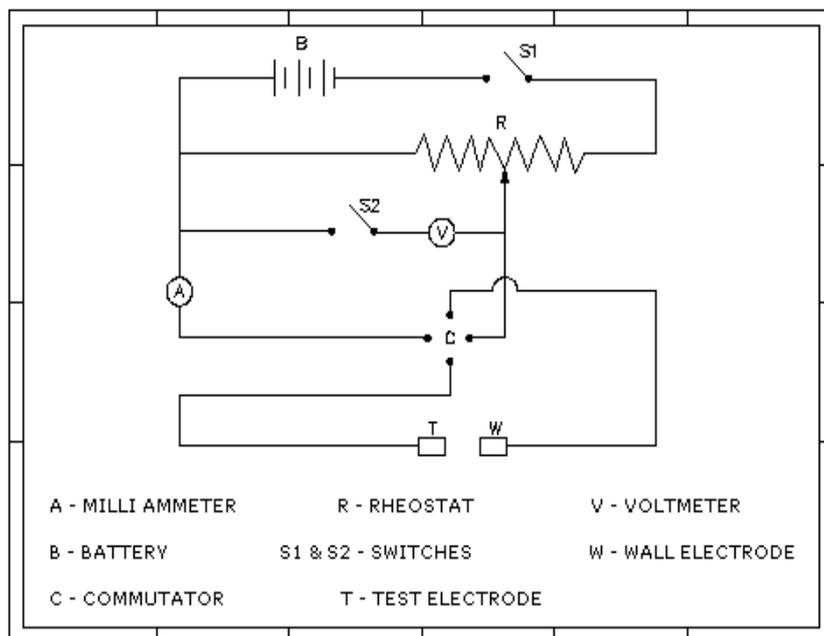
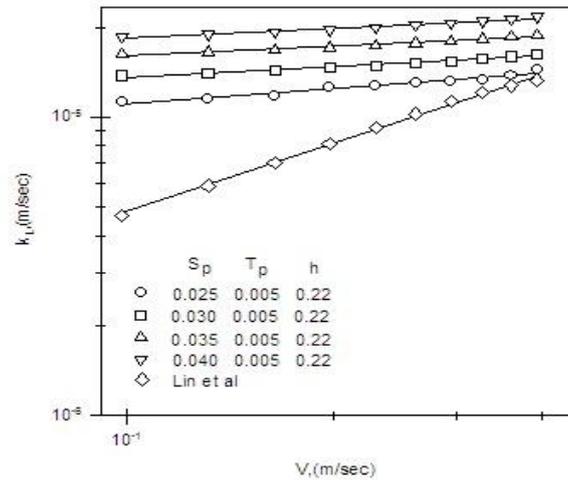
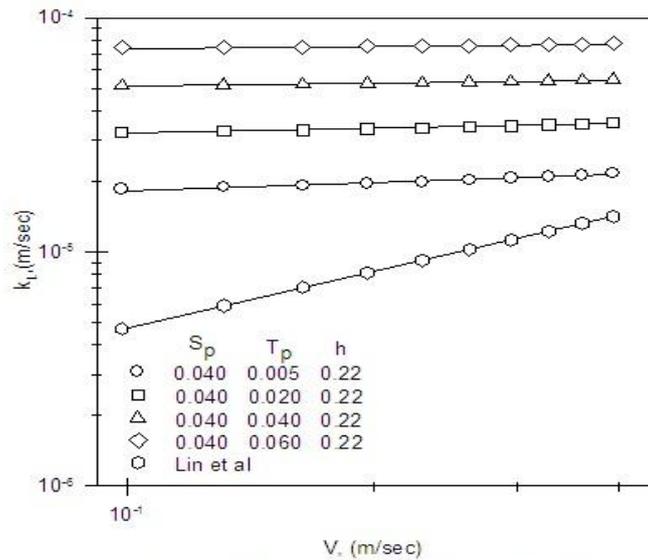


Figure 4: Circuit diagram



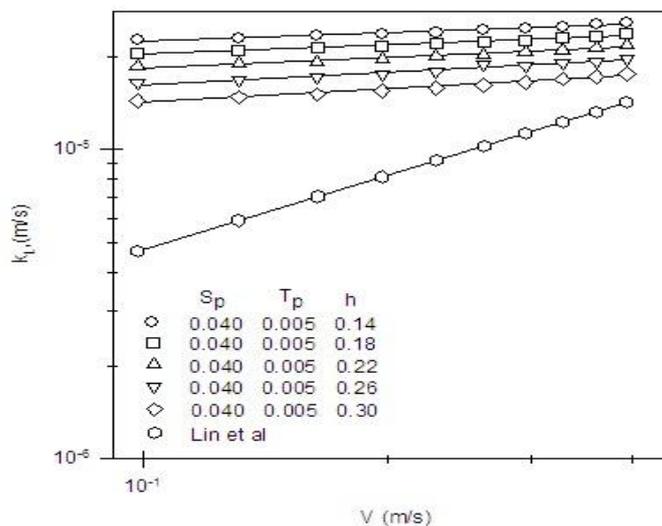
Variation of Mass Transfer Coefficient with Velocity
-Effect of Side of Triangular Plate

Figure 5



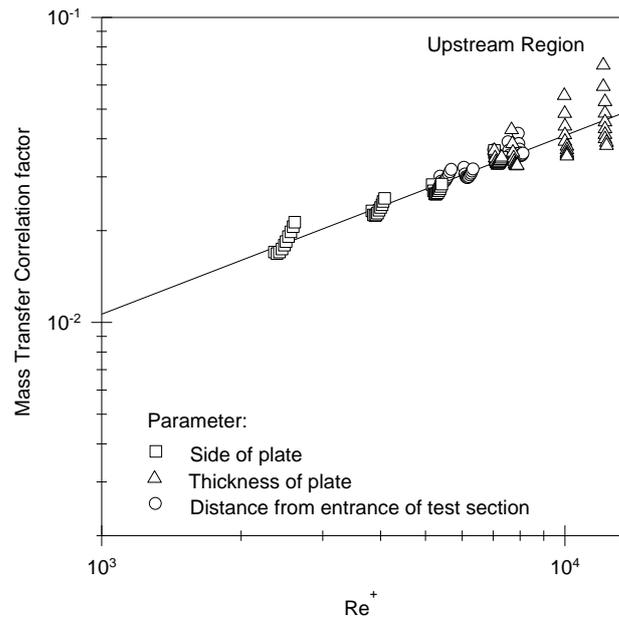
Variation of Mass transfer Coefficient with Velocity
- Effect of Thickness of Triangular Plate

Figure 6



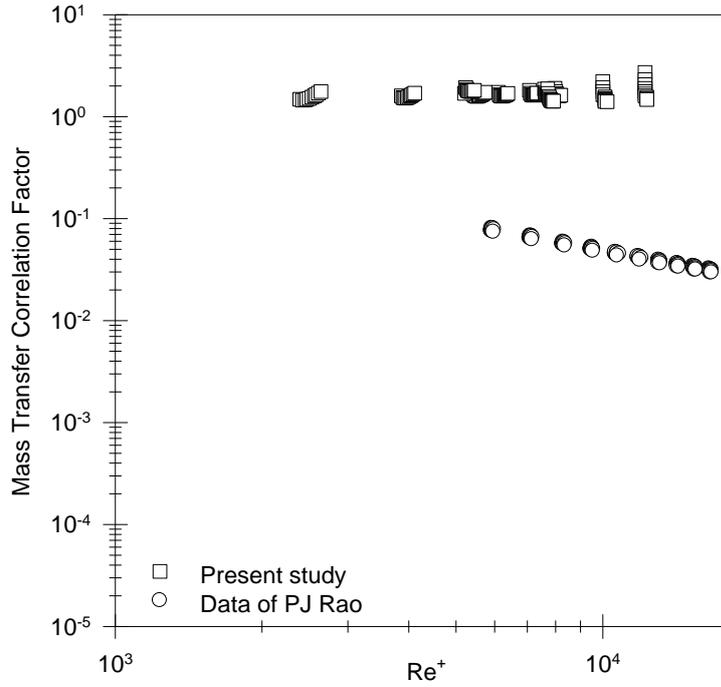
Variation of Mass Transfer Coefficient with velocity
- Effect of Distance of Triangular Plate from
entrance of Test Section

Figure 7



Correlation plot for Equation 9

Figure 8



Variation of CFR with Re^+

-Comparison plot

Figure 9

Analytical and Comparative Study on effect of Indian Classical Music on human body using EEG based signals

Ram K. Nawasalkar¹, Pradeep K. Butey²

¹Deptt. of Computer Science, Arts, Comm. and Science College, Kiran Nagar, Amravati.
S.G.B. Amravati University, Amravati, India

²Deptt. of Computer Science, Kamla Nehru Mahavidyalaya Nagpur.
R.T.M. Nagpur University, Nagpur, India

Abstract: Recently, Music is playing dominant role in human life. Due to heavy work load person can listens music to relax. The main aspect of the affective computing is to make computer more compatible to the user and this can be done by "getting into" users head to observe the mental state by using certain physiological device like EEG.

This paper describes now to recognize and percept emotions from brain signals while listening the Indian classical music measured with the electroencephalogram (EEG) device.

Keywords: Electroencephalography (EEG); Music; Brain activity; emotion; emotion recognition

I. Introduction

The music is dominant mood enhancer. Thus, most oftenly people listen a music since early in the morning till late night. Indian Classical Music is the soul of every music. Classical Music greatly affect on brain activity; it may have a positive effect on hormone system that's why people feel relax after hearing the classical music.

India has got the strong historical background of music. Archeological studies and evidence too has validated the presence of music from the ancient time. The 'Samaveda' includes hymns and describes the Indian music. While discussing about the Indian Classical Music, the striking word comes 'Raga'. It is the very heart of Hindustani Classical Music.

Shastra says every Raga impacts on human body and mind. A Raga is characterized by its own particular 'Ras' or 'Mood'. The acknowledged order of these nine sentiments, or emotions, is as follows: romantic and erotic, humorous, pathetic, anger, heroic, fearful, disgust, amazement and peaceful. Each Raga, in addition to being associated with a particular mood, is also closely connected to a particular time of day or a season of the year. Thus through the rich melodies and beat/ matra of Indian music, every human emotion, every subtle feeling and nature, can be musically expressed and experienced. In the ancient era it was difficult to prove the relation between emotion and music due to the lack of development in science and technology.

"Electroencephalography (EEG) is the recording of electrical activity along the scalp. EEG measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. The clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a short period of time, usually 20–40 minutes, as recorded from multiple electrodes placed on the scalp(hair bearing portion of the head). Diagnostic applications

generally focus on the spectral content of EEG, that is, the type of neural oscillations that can be observed in EEG signals. In neurology, the main diagnostic application of EEG is in the case of epilepsy, as epileptic activity can create clear abnormalities on a standard EEG study".[1]

It observed that while listening the music, brain parts are involved in processing music, this include the auditory cortex, frontal cortex, cerebral cortex and even the motor cortex [2].

It is an ancient, and very pervasive, idea that music expresses emotion. Apart from the copious literature to this effect contributed by composers, musicologists, and philosophers, there is also solid empirical evidence from psychological research, that listeners often agree rather strongly about what type of emotion is expressed in a particular piece.[3]

Most of us listen music of choice during leisure time or while working / studying. Music can be used as a tool to relieve tension/ stress, solitude, it also enhance the listeners mood [4]. These changes are reflected clearly in physiological system for human. Most of previous studies are mainly focused on effect of music on brain functions without any mental workload. In the present study, we designed an experimental paradigm to see the effect of music on brain functions rate based on experimenter's choice of like music (Jazz, carnatic and Hard rock) during with / without mental workload.[5]

Emotions are the most important living factor in human body. The need for computer applications which can detect the current emotional status of the user is ever growing [6].In an human communication a lot of research work is done for recognizing the emotions from face and voice with less accuracy.

Using EEG signals for emotion recognition provides a lot of advantages:

- Physiological signals are constantly emitted.
- As sensors are attached directly to the body, a person cannot move out of reach from a camera or a microphone that is placed in a room
- Biosignals are controlled by the central nervous system and therefore cannot be inuenced intentionally.

This research is limited to the those who are kin interested in Indian classical music.

II. Related Work

Emotions are part of any natural communication involving humans. They can be expressed either verbally through emotional vocabulary, or by expressing non-verbal cues such as intonation of voice, facial expressions and gestures.[7]

Music has significant effect on our body and mind.[8]

'Emotion in Motion' is an experiment designed to understand the emotional reaction of people to the variety of musical excerpts, via self-report questionnaires and the recording of electrodermal activity (EDA) and heart rate (HR) signals.[9]

'Emotion in Motion' is an experiment designed to understand the emotional reactions of people during music listening, through self-report questionnaires and the recording of physiological data using on-body sensors.[10]

III. Methodology

Electroencephalography Techniques

Similarity index(SI) is used to find the interdependency between different brain regions of musicians and non musicians while listening the music and in a condition of rest through EEG signals.

EEG recording was done for 180 seconds with a sampling rate of 128 HZ audio file to the person.

Both musicians and non musicians were asked to listen a raga by vocalist in the form of audio files. After we have find out EEG spectra in different areas of the brain cortex. An experimental study is needed while listening different ragas.

There are 16 persons in the range of 20 to 40 years of age. This has age range observed to have maximum stress in recent era. Who were having the musical knowledge or not, but kin interested in Indian classical music participated in the study and spectral power.

21 channel EEG were recorded simultaneously and the SP of each frequency band (alpha, beta, delta, theta) was calculated.

Emotion Acknowledgement (Recognition)

The acknowledgement of types of emotions depends on how the EEG features can be mapped on chosen emotion representation. The emotion representation used in the two dimensional mapping they are valence and arousal, which were already mentioned below

- Valence (Stimulation or brainwave): Valence, as used in psychology, especially in discussing emotions, means the intrinsic attractiveness (Positive Valence) or aversiveness (Negative Valence) of an event, object, or situation. However, the term is also used to characterize and categorize specific emotions. For example, the emotions popularly referred to as "Negative", such as anger and fear, have "Negative Valence". joy has "Positive Valence". positive, happy emotions result in a higher frontal coherence in alpha, while negative emotions result in higher right parietal beta power.

- Arousal (Excitement): One of the main function is to motivate new behaviors. The arousal that accompanies emotions has a significant impact on things such as anxiety, performance stress and ultimately performance. A higher beta power and coherence in the parietal lobe are represented excitation.

If valence will be hard to determine then arousal scale result can be used and the vice versa, both dimensions will have a good spread, and any emotion that can be mapped on the two axes can be recognized.

There are two methods of brain towards study of emotions of interest for representing them theoretically.

1. Method of darwin: basic emotions have evolved through natural selection. It proposed eight basic emotions: anger, fear, sadness, disgust, surprise, curiosity, acceptance, and joy.

2. Method of cognition: In this method emotions are mapped according to their valence and arousal.

Both methods of EEG are widely used, but the second representation is chosen for the purpose of research because of its simplicity and suitability. The following Fig. 1 shows the general indication of the positions of certain emotions based on the emotion labelling .i.e. active, negative, passive and positive. With eight basic emotions: angry, afraid, sad, depressed, calm, content, happy an excited.

Emotion Representation

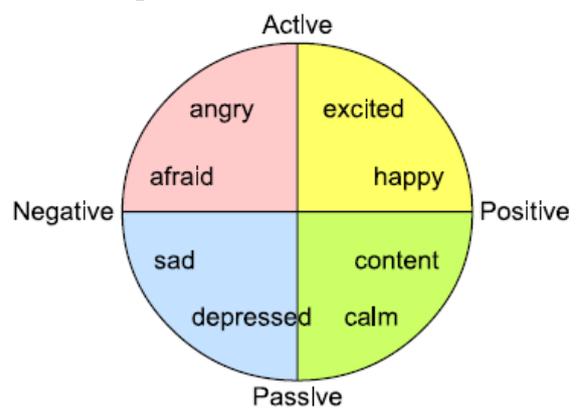


Figure 1: Arousal-valence model, labelled by Danny Qude Bos EEG-based Emotion Recognition[10]

Emotions in the Brain

Stimuli enter the brain at the brain stem (medulla oblongata). The limbic system which is like a cortical ring around the brain stem is responsible for initial emotional interpretation of signals from the autonomic nervous system. This part has also been found important for motivation and memory functions. Although motivation and memory also have their influence on the reaction to emotional stimuli; the rest of the text will focus on the limbic structures that are specifically relevant for emotional reactions.

Following are the steps of recognition of emotions from brain signals:

Step 1 To determine the optimal placement of a limited number of electrodes for any emotion recognizing device.

Step 2 putting this information into practice measured with the EEG device during the experimental phase.

Step 3- Finally analyze what are the results.

IV. Conclusion

In the present paper, the relationship between emotions and classical music are analyzed. An positive effect on brain after hearing of Indian classical music is more, as compared with other music after capturing the EEG signals.

The Indian classical music is found to be more effective on emotional status as compared to rock music in the state of quiet wakefulness.(open eyes)

Indian Classical Music can be used as a tool to relieve tension/ stress and to relax.

V. Future Scope

Future work will explore and investigate by means of similar relationship between emotions and different raga's as per the samay chakra.

We can also apply this method with ECG signals. We can apply this method to those people who are not aware of Indian Classical Music. We can apply this method on those people who are not having special musical knowledge.

References

1. <http://en.wikipedia.org/wiki/Electroencephalography> ite_note-Niedermeyer-1
2. Kristeva R, Chakarov V, Schulte-Monting J, Spreer J., Activation of cortical areas in music execution and imagining: a high-resolution EEG study, *NeuroImage* 2003 (20), pp: 1872 – 1883
3. Klaus R. Scherer and Marcel R. Zentner, Emotional Effects Of Music: Production Rules, Music and emotion: theory and research, *Oxford ; New York : Oxford University Press, Juslin, P.N. & Sloboda, J.A. (ed.)* (2001)
4. Lonsdale, A. J., & North, A. C., Why do we listen to music? A uses and gratifications analysis, *British journal of psychology London England* 1953,2011 pp: 102 (1), 108-134
5. B. Geethanjali, K. Adalarasu, R. Rajsekaran, Impact of Music on Brain Function during Mental Task using Electroencephalography, *World Academy of Science, Engineering and Technology* 66 2012
6. R. W. Picard and J. Klein, Toward computers that recognize and respond to user emotion: Theoretical and practical implications. *Interacting with Computers*, 14(2):141–169, 2002.
7. Chanel, Guillaume, Emotion assessment for affective computing based on brain and peripheral signals. *Thèse de doctorat : Univ. Genève*, 2009, no. Sc. 4123
8. Dr. K. Adalarasu , M. Jagannath, S. Naidu Keerthiga Ramesh, B. Geethanjali , A Review On Influence of Music on Brain Activity Using signal Processing and imaging System, *International Journal of Engineering Science and Technology (IJEST)*, Vol. 3 No. 4 Apr 2011
9. Javier Jaimovich¹, Niall Coghlan¹ and R. Benjamin Knapp², Emotion in Motion: A Study of Music and Affective Response, *9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)* 19-22 June 2012.
10. Danny Qude Bos EEG-based Emotion Recognition, *Department of Computer Science , University of Twente P.O. Box 217,7500 AE , Enschede*

Loss Reduction and Efficiency Improvement: A Critical Appraisal of Power Distribution Sector in India

Soham Ghosh

JRF, Department of Energy Science & Engineering, IIT Bombay, India

ABSTRACT: The main purpose of the present paper is to make an appraisal of the existing Power Distribution Sector in India with special focus on loss reduction and efficiency improvement of power supply. Different major aspects of technical and non-technical losses have been identified and on the basis of that a number of remedial measures have been suggested for loss reduction and to facilitate the improvement of overall efficiency of the power distribution system. This may provide further inputs to energy planners and managers.

Key words: Power Distribution, Loss of Power, Loss Reduction, Energy Efficiency Improvement.

I. INTRODUCTION

The power sector constitutes the backbone of the national economy of any country. Adequate electrical power with a high degree of reliability and quality is also the key to Indian economic growth. India is the 5th largest power producer in the world with the total power capacity of more than 1,45,000 MW. Despite growth in power generation capacity over the last 5-Year Plans, India is facing huge power deficit with peak power deficit of about 16%. Keeping in view the central position of the power sector for good quality of life and sustainable economic development, the Government of India has adopted the policy of providing "access to uninterrupted quality power supply at affordable costs to all by the year 2012. The responsibility of translating this vision into reality vests with the power sector and particularly the power distribution sector functionaries".

In last 20 years, the Indian power sector has witnessed tremendous growth both in size and capacity. In India, the current power installed capacity of power generation aims to increase it to 2,12,000 MW by the year 2012. However, in spite of such massive expansion, the power sector in India has not been able to match the rapidly growing demand for reliable and cost effective supply. Demand for power has continued to grow at a compound annual rate of 8% and has completely outstripped the supply leading to an ever widening gap. The power sector faces many challenges today in its march towards meeting its goal of "Power to All".

In the overall power development scenario in India, the transmission and distribution system constitutes the essential link between the power generating sources and the ultimate consumption of that. The optimum utilization of the generated power is not possible without the help of an adequate and efficient transmission and distribution system. In India, though the expansion of transmission systems has been carried out in a planned way based on detailed technical studies, however, the distribution system has grown in an unplanned and haphazard manner to meet the immediate

objective of meeting growing demands of consumers on an urgent basis. This approach, over the years, has created an inefficient distribution system contributing to very high Aggregate Technical and Commercial losses (AT & C losses) and poor quality with low reliability of power supply to consumers. Unfortunately, it has led to tremendous consumer dissatisfaction. It has also affected the financial performance of utilities. Thus, in the ongoing power sector reforms, the focus has rightly shifted to upgrading this Sub-transmission and distribution (ST & D) system and improving its efficiency to reduce AT & C losses.

II. TRANSMISSION AND DISTRIBUTION (T & D) LOSSES

In India, the fact is that all energy supplied to a distribution utility does not reach the end consumers. A substantial amount of energy is lost in the distribution system by way of technical losses. These inherent losses in transmission and distribution of electrical energy from the generating stations to the ultimate consumers should be reduced by eliminating or minimizing the causes of losses.

T & D loss is the difference between units injected into the system and the units billed to the ultimate consumers, which is generally expressed as percentage of units injected. It is generally calculated for a period of one financial year:

$$\text{Hence, T \& D losses (\%)} = \frac{[(\text{Energy input} - \text{Energy billed}) \times 100]}{\text{Energy input}} \quad \{\text{For a financial year}\}$$

The transmission and distribution (T & D) losses in our country, which were around 15% up to 1966-67, increased gradually to 23.28% by 1989-90. After a brief spell of reduction in T & D losses to 21.13% (1994-95), there has been an upswing and the losses reached a level of 33.98% during 2001-02. Since then, a reducing trend has been observed as T & D losses have come down to 32.54% during 2002-03, 32.53% during 2003-04 and 31.25% during 2004-05.

The Transmission and Distribution losses in advanced countries of the world have been ranging between 6 to 11%. Even in many developing countries, T & D losses are less than the level obtaining in India. However, T & D losses in India are not comparable with advanced countries as the system operating conditions there are different from those obtainable in India. As per the T & D losses issued by CEA, taking into consideration the Indian conditions, it would be reasonable to aim for containing T & D losses within 10 to 15% in different States (Table 1):

Table 1: Percentage of T & D Losses in States/ UTs of

India				
Sl. No	States/ UTs	2002-03	2003-04	2004-05
1	Haryana	37.65	32.07	32.11
2	Himachal	21.16	22.76	28.90
3	Jammu & Kashmir	45.55	45.54	41.08
4	Punjab	24.42	25.96	25.42
5	Rajasthan	42.61	43.74	44.68
6	Uttar Pradesh	34.16	35.17	34.39
7	Uttaranchal	25.17	49.23	39.30
8	Chandigarh	24.06	39.06	30.37
9	Delhi	45.82	43.66	45.40
10	BBMB	5.20	1.22	0.98
11	Gujarat	28.52	24.20	30.43
12	Madhya Pradesh	43.31	41.44	41.30
13	Chhattisgarh	37.86	42.55	28.06
14	Maharashtra	34.01	34.12	32.40
15	D & N Haveli	40.26	15.10	16.00
16	Goa	40.26	45.05	35.97
17	Daman & Diu	14.95	16.88	15.56
18	Andhra Pradesh	30.11	27.73	23.96
19	Karnataka	24.57	23.29	26.08
20	Kerala	27.45	21.63	22.48
21	Tamilnadu	17.31	17.16	19.28
22	Lakshadweep	11.29	11.85	10.20
23	Pondicherry	21.10	11.60	18.15
24	Bihar	37.98	36.66	38.88
25	Jharkhand	21.19	25.35	19.62
26	Orissa	45.36	57.09	44.02
27	Sikkim	54.85	54.99	50.49
28	West Bengal	25.93	31.01	28.54
29	A & N Islands.	19.78	25.95	12.63
30	DVC	3.34	2.69	2.69
31	Assam	38.30	39.31	51.76
32	Manipur	63.66	65.18	70.61
33	Meghalaya	21.92	16.73	28.35
34	Nagaland	56.71	55.00	48.26
35	Tripura	40.64	46.44	59.54
36	Arunachal Pradesh	38.95	47.54	42.96
37	Mizoram	46.91	55.54	66.14
	All India	32.54	32.53	31.25

Technical Losses

Technical loss is inherent in electrical systems, as all electrical devices have some resistance and the flow of current causes a power loss (I^2R loss). Integration of this power loss over time, i.e., $I^2R.dt$ is the energy loss. The summary of different types of technical losses is given in Table 2:

Table 2: Losses due to Technical Reasons

<ul style="list-style-type: none"> Line loss 	<ul style="list-style-type: none"> Loss in conductors/ cables where lower size conductors are used. This causes sags and temperature rise in conductors which further aggravate the loss, Loss in higher loaded phase wires due to unbalanced loading, Losses due to current in neutral for cases of unbalanced where neutral wires of lower size are used (like 3 ½ core cables, and neutral wires of size lower than phase wires), Loosening of strands (in multi-strand conductors like ACSR, AAC, AAA, etc.).
<ul style="list-style-type: none"> Losses in mid-span joints (or any joint) at terminations 	<ul style="list-style-type: none"> Contacts of joints due to improper installation and looseness, Contacts of joints due to inadequate surface area of contact.
<ul style="list-style-type: none"> Losses in transformers (typically DTs) 	<ul style="list-style-type: none"> Loose connections at brushings, Bend in jumpers at connectors where the strands are not tightly held, High no-load loss depending on type of core used, High no-load loss in repaired transformers, where the core has not been properly tightened, No-load loss in case a large number of lightly loaded DTs, High copper loss for transformers operating at sub-optimal loading which is not commensurate with the designed optimal loading.
<ul style="list-style-type: none"> Losses in service cables and connections 	<ul style="list-style-type: none"> Under sized service cables, Loss in joints of service cables at the poles or junction boxes, Use of inappropriate fasteners without spring washer at the crimped joints.

<ul style="list-style-type: none"> Loss due to high impedance faults 	<ul style="list-style-type: none"> Tree touching, creepers, bird nesting, Insulator breakages and tracking on the surface of the insulator.
<ul style="list-style-type: none"> Losses in re-wired fuses/ jumpers 	<ul style="list-style-type: none"> Loose connection, Inadequate size of fuse wires – often a source of hot spots.

Commercial Losses

Commercial losses are caused by non-technical or commercial factors namely pilferage, theft, defective meters, errors in meter reading, estimating un-metered supply of energy etc. The summary is given in Table 3:

Table 3: Commercial Losses

<ul style="list-style-type: none"> Loss at consumer end meters 	<ul style="list-style-type: none"> Poor accuracy of meters, Large error in capital CTs / PTs, Voltage drop in PT cables, Loose connection in PT wire terminations, Overburdened CT.
<ul style="list-style-type: none"> Tampering / bypass of meters 	<ul style="list-style-type: none"> Where meter without tamper-proof-tamper-deterrent/tamper-evident meters are used, Poor quality sealing of meters, Lack of seal issue, seal monitoring and management system, Shabby installation of meters and metering systems, Exposed CTs/ PTs where such devices are not properly securitized.
<ul style="list-style-type: none"> Pilferage of energy 	<ul style="list-style-type: none"> From overhead ‘bare’ conductors, From open junction boxes (in cabled systems), Exposed connection/joints in service cables, Bypassing the neutral wires in meters.
<ul style="list-style-type: none"> Energy accounting system 	<ul style="list-style-type: none"> Lack of proper instrumentation (metering) in feeders and DTs for carrying out energy audits, Not using meters with appropriate data login features in feeders and DT meters, Lack of a system for carrying out regular

	<ul style="list-style-type: none"> (monthly) energy accounting to monitor losses, Errors in sending end meters, CTs and PTs, Losses connections in PT wires (which results in low voltage at feeder meter terminals), Energy accounting errors (by not following a scientific method for energy audits).
<ul style="list-style-type: none"> Errors in meter reading 	<ul style="list-style-type: none"> Avoiding meter reading due to several causes like house locked, meter not traceable, etc. Manual (unintentional errors) in meter reading, Intentional errors in meter reading (collusion by meter readers), Coffee shop reading, Data punching errors by data entry operators, Lack of validation checks, Lack of management summaries and exception reports on meter reading.
<ul style="list-style-type: none"> Error in bills 	<ul style="list-style-type: none"> Errors in raising the correct bill, Manipulation/ changes made in meter reading at billing centres – lack of a system to assure integrity in data, Lack of system to ensure that bills are delivered.
<ul style="list-style-type: none"> Receipt of payment 	<ul style="list-style-type: none"> Lack of system to trace defaulters including regular defaulters, Lack of system for timely disconnection, Care to be taken for reliable disconnection of supply (where to disconnect).

Revenue Loss due to Loss of Opportunity to serve

The revenue loss due to ‘loss of opportunity to serve’ is a very relevant but less visible aspect of revenue loss. Guarding against technical and non-technical losses is one aspect, but there is an equal need to guard against loss of revenue due to lost opportunity (Table 4):

Table 4: Reasons for Loss of Opportunity to Serve

<ul style="list-style-type: none"> System outage (any part or total systems) leading to loss of revenue 	<ul style="list-style-type: none"> Due to break down, Due to preventive maintenance, Due to load shading.
<ul style="list-style-type: none"> Overloading and unbalancing 	<ul style="list-style-type: none"> High voltage drop in lines leading to low voltage at consumer premises and lesser consumption, High voltage drop at tail end of affected phases in case of unbalance, leading to lesser consumption, High voltage drop due to large reactive currents causing I²R drop in lines and resulting in lesser voltage and low consumptions.
<ul style="list-style-type: none"> Tree touching 	<ul style="list-style-type: none"> Voltage sags at consumer premises, leading to lower power consumption.

Aggregate Technical and Commercial Loss

The aggregate of T & D loss and loss due to non-realization of billed demand is termed as aggregate technical and commercial loss (AT & C loss).

$$AT \& C \text{ loss } (\%) = [(Energy \text{ input} - Energy \text{ realized}) \times 100] / Energy \text{ input}$$

$$Energy \text{ realized} = [Energy \text{ billed} \times Collection \text{ efficiency}]$$

$$Collection \text{ efficiency } (\%) = [Amount \text{ Realized} \times 100] / Amount \text{ billed}$$

AT & C loss is a transparent measure of the overall efficiency of the distribution business as it measures technical as well as commercial losses. The schematic diagram (Fig.1) shown below captures the essential components of AT & C losses and is translated in terms of units both at the physical level of transmission and distribution and at the financial level of billing and collection:

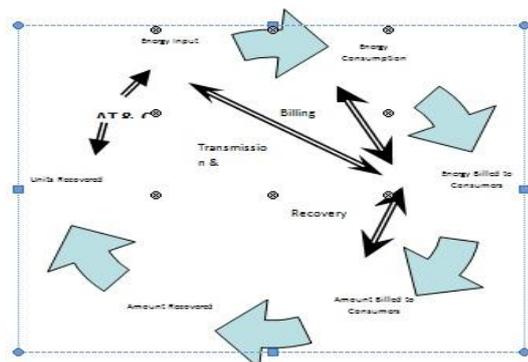


Fig.1: Schematic Diagram Showing Components of AT & C Losses

The AT & C losses are presently in the range of 18 % to 62% in various States. The national average AT & C loss of the distribution companies in the year 2002-03 was 36.63% and has reduced to 33.82% in the year 2004-05, though there is a wide variation of losses among the States and variation among Districts within the State.

III. SCOPE OF LOSS REDUCTION

Technical Loss Reduction

The technical losses in our power systems are high especially in distribution systems. The causes of high technical losses are varied and require different remedial measures to be implemented to bring them down to acceptable levels.

Short-term Measures

The immediate improvement and reduction of losses in the technical system are based upon sample studies, statistical scrutiny/ analysis of the information/ data collected from the existing system detail. These are:

- i. Network Reconfiguration – It gives an option to handle the increased demand and increases system reliability. It is effective when voltage drops between the nodes to be linked is rich and the distance between the nodes is short. Within a feeder it is effective only when the zigzag factor is high.
- ii. Network Reconductoring – The size of conductor/ cable determines the current density and the resistance of the line. A lower conductor size can cause high I²R losses and high voltage drop which causes a loss of revenue as consumer’s consumption and hence revenue is reduced. The recommended practice is to find out whether the conductor is able to deliver the peak demand of the consumers at the correct voltages, that is, the voltage drop must remain within the allowed limits specified in Electricity Act, 2003.
- iii. Preventing Leakages at Insulators - Cracking of insulator and flashover across insulators often cause outages and result in loss of revenue. Use of appropriate material for insulators, depending on the nature of pollution, and designed protected creepage path helps in reducing insulator failure. Preventive actions are regular inspection and hot line washing.
- iv. Automatic Voltage booster – it is similar to that of the series capacitor as an on-load tap changer it boosts the voltage at its point of location in discrete steps. This, in turn, improves the voltage profile and reduces the losses in the section beyond its point of location towards the receiving end. It has a total voltage boosts of 10% in four equal steps and the loss reduction is directly proportional to voltage boosts.
- v. Better Management of Distribution Transformers – the following measures can be taken in this regard:
 - a. Augmentation/ addition of distribution transformers ;
 - b. Relocation of distribution transformers at load centers;
 - c. Low voltage (less than declared voltage) appearing at transformers consumers terminals;
 - d. Guarding against loss in transformers through oversized transformers operating at low loading, undersized transformers, unbalanced loads in secondary side, connector at bushings, low oil level/ oil leakages, hot spots in core, use of energy efficient transformers etc.

- vi. Load Balancing and Load Management – if the loads on each of the three phases of a distribution lines or among feeders are redistributed, the losses will be reduced. The best method to identify load balance is to construct current duration curves for all three phases. In the scenario of overloaded distribution systems, load management plays a very important role for reduction of technical losses. Distribution automation along with SCADA (Supervisory Control and Data Acquisition System) is an important tool for load management which should be introduced.
- vii. Capacitor Installation – the use of capacitors to correct for poor power factor is a well established and cost effective means of reducing distribution system losses and maximizing the revenue. In most LT distribution circuits, it is found that the power factor (PF) ranges from 0.65 to 0.75. For low PF the amount of current drawn increases to meet the same kW demands of load. Overall improvement in the operating condition can be brought about by reducing the system reactance. This can be done by the application of shunt capacitor in the following ways – across individual customers, advantage points on LT and 11 kV feeders, at distribution transformers and at 33/11 kV sub stations.
- viii. Improving joints and connections – Improper joints are a source of energy loss in both overhead and underground systems. The conductivity of joints should not be less than an equivalent length of the conductor. Joints should be capable of carrying maximum fault current without failure or deterioration for the time required for the protective system to operate.
- ix. Increase in HT/LT ratio – It is well known that for high HT/ LT ratio, the losses will be low. The losses for a given quantum of power supplied by a line are inversely proportional to the square of its operating voltage. Higher the operating voltage, lower will be the line losses. Therefore, by increasing the HT lines the losses will be reduced.
- x. Adoption of high voltage distribution system (HVDS) – Adoption of HVDS by converting existing LVDS to HVDS reduces the technical losses appropriately.
- xi. Preventive and regular maintenance – These components of the distribution system are necessary to reduce/ eliminate breakdowns. Care should be taken to optimize preventive maintenance, because each shutdown due to preventive maintenance is also a source of revenue loss. It can be minimized by careful design and healthy installation practices. The following activities should be undertaken for preventive maintenance:
 - a. Maintenance of overhead lines,
 - b. Correction of bent poles,
 - c. Rewinding transformer,
 - d. Monitoring transformer tank temperature,
 - e. Use of protective devices,
 - f. Improved bushings,
 - g. Transformer oil testing,
 - h. Repairing of broken parts.

Long-term Measures

Long-term measures for technical loss reduction involve all measures that need to be taken for the improvement of quality and reliability of power supply and reduction of

T & C losses in a given area. These include upgrading, strengthening and improvement of the sub-transmission and distribution system in a circle to meet the future load demand for the next five years. The activities for preparation of a long term plan are listed below:

- i. Data collection regarding existing loads, operating conditions, forecast of expected loads, etc. from grid sub-station upto consumers level;
- ii. Mapping of existing system;
- iii. Analysis of existing system;
- iv. Load forecast;
- v. Plan for upgrading the network;
- vi. Technology options including integration of features for modernization of system;
- vii. Evaluation of various alternatives for least cost optimal solution;
- viii. Firming up of scope of works;
- ix. Preparation of cost estimation;
- x. Phasing of works and their cost;
- xi. Financial analysis.

Acceptable Technical loss levels

Acceptable technical loss levels depend on economic factors such as cost of power and energy, costs of equipment and discount rates rather than purely on technical factors. The achievable level of losses is subject to various factors given below in Table 5:

Table 5: proposed Targets for economic loss levels

S. No.	System Component	Levels for Peak Power losses	
		Target Level %	Max Tolerable %
1.	Step up transformer and EHV transmission system	0.50	1.00
2.	Transmission to intermediate voltage level, transmission system and step-down to sub-transmission voltage level	1.50	3.00
3.	Sub-transmission system and step-down to distribution voltage level	2.25	4.50
4.	Distribution lines and service connections	4.00	7.00
	Total power losses	8.25	15.50

Commercial Loss Reduction

Almost all the commercial losses occur at the distribution stage and that is where action has to be taken to control that.

Reasons for commercial losses

The major reasons for commercial losses are:

- i. Direct tapping by non-customers – unscrupulous consumers extract electricity illegally by bypassing the energy meter or by connecting leads directly to distribution lines. This kind of power theft takes place

- mainly in domestic and agricultural sectors. This should be tackled on a priority basis by the utility.
- ii. Pilferage and theft of energy by existing customers – theft of electrical energy by existing customers is causing an increase in revenue losses. Emphasis can be given on inspecting high value services for more effective and immediate gains.
 - iii. Defective metering, billing and collection – commercial losses are also caused by some deficiencies in commercial function of the utility, viz., metering, billing and collection. Though these losses are not due to any deliberate action of the customers, they are due to internal short coming of utility which can be tackled easily.
- b. Having clearly visible and accessible seals that can be subjected to easy inspection;
 - c. Mounting the meter and CTs inside a box with a clear window;
 - d. Ensuring height and location of installation for easy readability of meters;
 - e. Locating meters in public domain in full public view;
 - v. Measures for improvement in billing and collection – Correct billing and timely delivery of bills go a long way in improving the revenue collections. The normal complains viz. non-receipt / late-receipt of bills, wrong bills, wrong reading status, wrong calculations etc. should be avoided.
 - vi. Users' Associations, Panchyats and Franchisees in Billing and Collection – The electricity Act, 2003, visualizes the role of users' associations, co-operatives, panchayats and franchisees in electricity distribution management be successfully inculcated to develop a sense of belonging to and stake in the entity.
 - vii. Legal Measures for Reducing Commercial Losses – The Electricity Act, 2003, has brought radical changes in all the facets of the electricity sector. The relevant Sections of the Act are 55, 126, 127, 135, 138, 145, 150, 151, 152, 153, 154, 156, 157, 168, 169, 170 and 171.

Measures for commercial loss reduction

The measures for reducing commercial losses depend on the factors that cause them. These are:

- i. Measures for controlling direct tapping by non-customers and customers –
 - a. Stopping theft by direct tapping;
 - b. Use of aerial bunched cables/ partial insulated LT lines;
 - c. Public relation and awareness campaigns by utility;
- ii. Measures for controlling pilferage of energy by existing customers –
 - a. The energy meter should be housed in a separate box sealed and made inaccessible to the consumers. The fuse cut-outs should be provided after the meter;
 - b. Multi-core PVC cables should be used as service mains instead of single core wires,
 - c. Severe penalties may be imposed for tampering with metering seals, etc.;
 - d. Theft of electricity should be publicized as a social and economic crime and people should be informed of the provisions in electricity laws in this regard.
- iii. Measures for reducing defective metering –
 - a. Stuck up meters;
 - b. No reading furnished by the meter reader, for a good no. of services, at times continuously;
 - c. Constant nil consumption cases reported without any comment;
 - d. Progressive readings recorded in disconnected services;
 - e. No relation between the meter capacity and the load;
 - f. Adoption of wrong multiplication factors (MF) for billing as the change in MF in not intimated to the billing agency.
- iv. Meter installation – It is often considered as a low skill, labor oriented activity. It must be given due importance to against revenue loss. Certain installation practices to prevent this are:
 - a. Having a visually traceable and joint free incoming cable, shrink wrapped sealed joint;

IV. CONCLUSION

From the above discussions it is found that there are various factors responsible for AT & C losses which need to be eliminated. The approaches taken over the years in India has created an inefficient distribution system contributing to very high AT & C losses and poor quality and reliability of power supply to consumers. It has led tremendous consumes dissatisfaction as well as it has affected the financial performance of the utilities. As it is extremely difficult to eliminate all the causes simultaneously in our country, strategically measures should be taken to reduce or marginalize the major causes of losses. In the ongoing power sector reforms, the focus has rightly been shifted to upgrading the sub-transmission and distribution (ST & D) system and improving its efficiency to reduce AT & C losses. Ultimately, this may contribute in the process of overall national development.

References

- [1] GoI, Planning Commission, Eleventh Five-Year Plan (2007-12), Power & Energy: Energy Policy & Rural Energy, New Delhi, 2007.
- [2] IGNOU, School of Engineering and Technology, Power Distribution Sector (BEE-001), New Delhi, 2007.
- [3] GoI, Ministry of Power, Policy Document on Electricity for All, New Delhi, 2002.

Boundary Layer Flow and Heat Transfer Flow past Stretching Sheet with Emperature Gradient Dependent Heat Sink and Internal Heat Eneration

J. Phakirappa, P. H. Veena, V. K. Pravin

*Dept. of Mathematics, Vijayanagar College of Engg. Bellary, Karnataka, India
 Dept. of Mathematics, Smt.V.G.College for Women Gulbarga, Karnataka, India
 Dept. of Mechanical Engg, P.D.A. College of Engg. Gulbarga, Karnataka, India*

ABSTRACT

Free convection viscous fluid flow and heat transfer flow past a stretching sheet with temperature gradient dependent heat sink and internal heat generation is considered for the study. The exact solutions of momentum equation and energy equation are obtained in terms of Kummer's function and the effects of various parameters like magnetic parameter (Mn), permeability parameter (k_2) and suction parameter (S) are analyzed by drawing various graphs for tangential and vertical velocity and temperature field.

NOMENCLATURE

<p>c stretching rate</p> <p>C_p specific heat</p> <p>E elastics parameter</p> <p>G dimensionless stream function</p> <p>k- thermal conductivity</p> <p>k_0 coefficient of viscosity</p> <p>k_1 visco-elastic parameter</p> <p>k_2 permeability parameter</p> <p>Mn magnetic parameter</p> <p>Pr prandtl number</p> <p>Q heat source/sink parameter</p> <p>q_w local wall heat flux</p> <p>S suction parameter</p> <p>T temperature</p> <p>T_w temperature at wall</p> <p>T_∞ temperature away from wall</p> <p>u,v velocity components along x- and y- directions</p> <p>v_0 suction velocity</p>	<p style="text-align: center;">Greek symbols</p> <p>γ coefficient of kinematic viscosity</p> <p>ρ fluid density</p> <p>σ electrical conductivity</p> <p>β heat generation parameter</p> <p>η dimensionless similarity variable</p> <p>ξ dimensionless transformed variable</p> <p>θ dimensionless concentration variable</p> <p>ε small variable</p> <p style="text-align: center;">Subscripts</p> <p>w condition at the wall</p> <p>∞ condition at infinity</p> <p>η derivative with respect to η</p>
--	---

I. INTRODUCTION

The study of boundary layer behaviours over a stretching sheet occurring in several engineering applications and manufacturing processes in industry. The practical applications of continuous flat surfaces are in aerodynamic extrusion of plastic sheets, rolling and manufacturing of plastic films, cooling of metallic plates and boundary layer flow over heat treated materials between feed roll and a windup roll.

Sakiadis [1] initiated the study of boundary layer over a continuous solid surface, flat surface and the cylindrical surfaces, Mc-Carmack and Crane [2] presented an analysis on boundary layer flow caused by stretching of elastic flat surfaces and between two surfaces under various physical situations.

The investigations have a definite bearing on the problem of polymer sheet extruded continuously from a dye. It is generally assumed that the sheet is inextensible. But many cases arises in polymer industry in which it is necessary to deal with a stretching sheet as noted by Crane [3]. Gupta and Gupta [4] presented the heat and mass transfer over a stretching sheet with blowing or suction. Grubka and Bobba [5] have worked on the heat transfer occurring on a linearly stretching surface under variable temperature. Bujurke et. al. [6] made an investigation on the heat transfer analysis in a second order fluid flow past a stretching surface with heat transfer. Datta et. al. [7] have studied the distribution of temperature in a continuous stretching sheet with uniform wall heat flux. Chen and Char [8] analyzed the heat transfer on continuous stretching sheet with two different cases viz power law surface temperature (PST) and power law surface heat flux (PHF). Further flow and heat transfer from a linearly stretching sheet gained more importance due to practical applications in industrial processes. Abel and Veena [9] have analyzed visco-elastic fluid flow and heat transfer in a porous medium over a stretching sheet, Vajravelu and Rollins [10] studied heat transfer properties in a visco-elastic fluid over a continuous stretching sheet with power law surface temperature (PST) and power law surface heat flux (PHF) by considering heat transfer characteristics in a fluid initially at rest and at a uniform temperature. Further Vajravelu and Nayefh [11] presented

the flow and heat transfer by introducing temperature dependent heat source or sink. They considered heat transfer in a saturated porous medium over a continuous impermeable stretching surface with power law surface temperature (PST) and power law surface heat flux (PHF) including the effects of fractional heat and internal heat generation or absorption.

Many authors including Anjalidevi and Thiyagaraj [12], Ranjagopal et. al. [13], Sanyal and Das gupta [14], Sujit Kumar Khan et. al. [15], Takhar and Soundalgekar [16], Mahapatra et. al. [17], Bhargava et. al.[18], Idress and Abel [19], Takhar et. al.[20], have analyzed the problem on boundary layer flow due to the stretching sheet/continuous moving sheet for different flow models and boundary conditions.

Motivated by these analyses, in the present paper we studied the free convection flow past a non-isothermal stretching surface in the presence of porous medium and temperature gradient dependent heat sink including the internal heat generation. The exact solutions of momentum equation and energy equation in terms of Kummer's function are developed. The result of the study are discussed for different numerical values of the parameters like magnetic parameter (Mn), permeability parameter (k_2), suction parameter (S), Prandtl number (Pr) through graphs.

II. MATHEMATICAL FORMULATION

We consider a steady two dimensional boundary layer free convection viscous fluid flow past a stretching sheet in porous medium in the presence of temperature gradient dependent heat sink for the study.

The flow is produced due to stretching of the sheet by applying two equal and opposite forces along x-axis, keeping origin fixed. The flow is assumed in a region $y > 0$, in order to get the effect of temperature difference between the surface and the ambient fluid. Hence the governing boundary layer equations for momentum and heat transfer are of the following form.

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad \dots 1$$

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \gamma \frac{\partial^2 u}{\partial y^2} - k_0 \left[u \frac{\partial^3 u}{\partial x \partial y^2} + v \frac{\partial^3 u}{\partial y^3} + \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial y^2} - \frac{\partial u}{\partial y} \frac{\partial^2 u}{\partial x \partial y} \right] - \frac{\sigma \beta_0^2 u}{\rho} - \frac{\gamma}{k^1} u \quad \dots 2$$

$$u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \frac{k}{\rho C_p} \frac{\partial^2 T}{\partial y^2} + Q^1 \frac{\partial T}{\partial y} + \frac{Q}{\rho C_p} (T - T_\infty) \quad \dots 3$$

Where u and v are the velocity components in x- and y- directions respectively. k-the thermal conductivity of the fluid medium. K_0 - coefficient of visco-elasticity,

Q-the heat source/sink parameter, γ -the coefficient of kinematic viscosity.

The appropriate boundary conditions for velocity are of the form.

$$\begin{aligned} u = c x \quad v = -v_0 \quad \text{at } y = 0 \\ u = 0 \quad \text{as } y \rightarrow \infty \end{aligned} \quad \dots 4$$

where c is a constant known as stretching rate and v_0 is the suction velocity.

III. VELOCITY TRANSFER ANALYSIS

To solve the equations (1) and (2) we assume a solution for velocity field as below

$$u = c x G'(\eta), \quad v = -\sqrt{c\gamma} G(\eta) \quad \text{and } \eta = \sqrt{\frac{c}{\gamma}} y \quad \dots 5$$

Obviously for the above u and v, continuity equation (1) is identically satisfied. Substituting (5) in momentum equation (2), it reduces to fourth order non-linear ordinary differential equation of the form

$$G'^2(\eta) - G(\eta)G''(\eta) = G''''(\eta) - k_1 [2G'(\eta)G'''(\eta) - G(\eta)G''''(\eta) - G''^2(\eta)] - (Mn + k_2)G'(\eta) \quad \dots 6$$

where $k_1 = \frac{k_0 c}{\gamma}$ - visco-elastic parameter

$Mn = \frac{\sigma \beta_0^2}{\rho c}$ - magnetic parameter

$k_2 = \frac{\gamma}{k^1 c}$ - permeability parameter

and the corresponding boundary conditions (4) reduces to

$$\begin{aligned} G'(\eta) = 1 \quad G(\eta) = -\frac{v_0}{\sqrt{c\gamma}} \quad \text{at } \eta = 0 \\ G'(\eta) = 0 \quad G''(\eta) = 0 \quad \text{as } \eta \rightarrow \infty \end{aligned} \quad \dots 7$$

Here $S = \frac{v_0}{\sqrt{c\gamma}}$ - Suction parameter.

The exact solution of momentum equation (6) subjected to the boundary conditions (7) is of the type,

$$G(\eta) = A + Be^{-E\eta} \quad \dots 8$$

where $A = \frac{E^2 - (k_2 + Mn)}{E}$

$$B = -\frac{1}{E}$$

and

$$E = \frac{1}{2} \left[S + \sqrt{S^2 + 4(Mn + k_2 + 1)} \right]$$

Then exact solution equation (8) can be expressed as

$$G(\eta) = \frac{1}{E} [E^2 - (k_2 + Mn) - e^{-E\eta}] \quad \dots 9$$

Hence we have

$$G'(\eta) = e^{-E\eta} \quad \dots 10$$

IV. HEAT TRANSFER ANALYSIS

In the energy equation (3) where k - the thermal conductivity varies approximately with temperature and the temperature gradient dependent heat sink is a linear function of the temperature. The appropriate boundary conditions for heat transfer boundary flow are

$$T = T_w = T_\infty + b x^p \quad \text{at } y = 0 \quad \dots 11$$

and $T = T_\infty$ as $y \rightarrow \infty$

where $T = T_w$ and $T = T_\infty$ are the temperature at wall and the temperature far away from the wall. b is a constant value which depends upon the property of the fluid.

To solve energy equation (3) we introduce the following non-dimensional quantity for temperature T as

$$\theta = \frac{T - T_\infty}{T_w - T_\infty} \quad \dots 12$$

and expression for temperature T as

$$T = T_\infty + b x^p \theta(\eta) \quad \dots 13$$

By using the transformation given in equations (4), (12) and (13), energy equation (3) reduces to the following form.

$$\theta'' + Pr G(1 + Q^*)\theta' - Pr(G' + \beta)\theta = 0 \quad \dots 14$$

Where $Q^* = -\sqrt{\frac{C}{\gamma}} Q$ - volumetric rate of heat generation

$$Pr = \frac{\mu C_p}{k} \text{ - Prandtl number}$$

$$\beta = \frac{Q}{\rho c C_p} \text{ - heat generation parameter}$$

and the corresponding boundary conditions are

$$\theta(\eta) = 1 \text{ at } \eta = 0 \quad \dots 15$$

$$\theta(\eta) \rightarrow 0 \text{ as } \eta \rightarrow \infty$$

To obtain the solution of equation (14) we define a new change of variable ξ as

$$\xi = -\frac{Pr(1 + Q^*)e^{-E\eta}}{E^2} \quad \dots 16$$

By using equations (13) and (16), equation (14) reduce to the form

$$\xi \theta_{\xi\xi} + \left[1 - Pr(1 + Q^*) \left(1 - \frac{k^{-1}}{E^2} \right) - \xi \right] \theta_\xi - Pr \left[\frac{\beta}{E^2 \xi} - \frac{1}{Pr(1 + Q^*)} \right] \theta = 0 \quad \dots 17$$

and the corresponding conditions (15) converts to

$$\theta \left(\xi = -\frac{Pr(1 + Q^*)}{E^2} \right) = 1, \quad \theta(\xi = 0) = 0 \quad \dots 18$$

The solution of the equation (17) subjected to the boundary conditions (18) is obtained and expressed in the following form of confluent hyper-geometric function (i.e. Kummer's function) of the similarity variable η as

$$\theta(\eta) = e^{-E\delta_1\eta} \frac{M(2S_1 - k_1, 1 + b_o, -k_2 e^{-E\eta})}{M(2\delta_1 - k, 1 + b_o, -k_2)} \quad \dots 19$$

Where $\delta_1 = \frac{a_o + b_o}{2}$

$$a_o = Pr(1 + Q^*) \left(1 - \frac{k^{-1}}{E^2} \right)$$

$$b_0 = \frac{1}{E} \sqrt{\text{Pr}^2 (1 + Q^*)^2 (E^2 - k^{-1})^2 + 4 \text{Pr} p \beta}$$

$$k_1 = \frac{p}{1 + Q^*} \quad \text{and} \quad k_2 = -\frac{\text{Pr}(1 + Q^*)}{E^2}$$

Non-dimensional wall temperature gradient from equation (19) is derived as

$$\theta'(0) = -E \delta_1 M (2\delta_1 - k_1, 1 + b_0, -k_2) + \frac{\left(\frac{2\delta_1 - k_1}{1 + b_0} \right) M (2\delta_1 - k_1 + 1, 2 + b_0, -k_2)}{M (2\delta_1 - k_1, 1 + b_0, -k_2)} \quad \dots 20$$

The local wall heat flux can be expressed as

$$q_w = -k \left(\frac{\partial T}{\partial y} \right)_{\eta=0} = -k b x^p \sqrt{\frac{c}{v}} \theta'(0) \quad \dots 21$$

SKIN FRICTION

After velocity transfer skin friction co-efficient in non-dimensional form is obtained as

$$T_o = -\mu \left(\frac{\partial u}{\partial y} \right)_{\eta=0} = u_o x, \quad \text{where } u_o = \mu \beta E^2 e^{-E\eta} \quad \dots 22$$

V. RESULTS AND DISCUSSION

The equations for free convection of incompressible viscous fluid flow and heat transfer past a stretching sheet with temperature gradient dependent heat sink and internal heat generation are examined. Energy and momentum equations are solved analytically and expressed in terms of Kummer's function. The effect of various physical parameters like Magnetic parameter (Mn), Permeability (k_2), Suction parameter (S) are examined on velocity profiles and temperature profiles. The value of Prandtl number is considered to be $\text{Pr} = 0.7$ which refers to air. Variation in longitudinal and transverse velocities are depicted for different values of Mn, k_2 and S. The longitudinal velocity is calculated for S at $\text{Mn} = 25$, $k_2 = 5$ and $c \times x = 1.0$. The transverse velocity is calculated for various values of S at $\text{Mn} = 25, k_2 = 5$ and $\sqrt{c\gamma} = -1.0$.

Fig.1(a) shows the variation of u versus η for various values of suction parameter $S = 0, 0.5, 1.0, 1.5$. We observed from the figure that there is a steady decrease in u with increase in S . The longitudinal velocity is maximum at the wall for all cases.

Fig.1(b) shows the graph of transverse velocity v versus η for $S = 0, 0.5, 1.0, 1.5$. We observed that there is steady increase in v with increase in η from $\eta = 0$ to $\eta = 0.7$. The transverse velocity steadily increases with the increase in the value of η .

Fig.2(a) is depicted to study the variation of $G(\eta)$ versus η for different values of visco-elastic parameter (k_1). It is noticed from the figure that $G(\eta)$ increases with increase of visco-elastic parameter (k_1).

Fig.2(b) presents the variation of the function $G(\eta)$ versus η for different values of magnetic parameter (Mn). From the figure it is observed that $G(\eta)$ increases with increase in the value of magnetic parameter (Mn).

In fig.2 (c) the variation of the function $G(\eta)$ versus η for different values of Suction parameter (S) is shown. From the figure it follows that $G(\eta)$ increases with increase in Suction parameter (S).

Fig.3(a) is drawn to show the variation of temperature $\theta(\eta)$ versus η for different values of Mn along with $k_2 = 20$, $\text{Pr} = 0.7$, $\beta = 2$ and for different combinations of p , Q^* and S. We observe from the figure that the maximum temperature corresponds to the curve I for which $\text{Mn} = 20$, $p = 0.5$, $S = 0.5$ and $Q^* = 0.5$. The temperature lowered for the rise in Mn for different combinations of p , Q^* and S. Wall temperature parameter plays an important role for lowering the temperature. The boundary layer thickness decreases as the magnetic parameter (Mn) increases which results in high temperature distribution at the wall.

Fig.3(b) shows the variation of temperature $\theta(\eta)$ versus η for effect of different values of permeability parameter (k_2) along with $\text{Mn} = 20$, $\text{Pr} = 0.7$, $\beta = 2$ and for different combinations of p , Q^* and S. We observe from the figure that the temperature profile decreases on the boundary layer with increase in k_2 for different combinations of p , Q^* and S. The temperature is unchanged at the wall with the change in physical parameters.

Fig.4 (a) is depicted in regard of wall temperature gradient $\theta'(0)$ versus Prandtl number (Pr) for different values of k_2 , Mn and β with the combinations of different k values p , Q^* and S. It is observed from the figure that temperature gradient increases with increase of values of Prandtl number (Pr).

Fig.4(b) shows the variation of wall temperature gradient $\theta'(0)$ versus magnetic parameter (Mn) for different values of k_2 , Pr and β with different combinations values of p , Q^* and S. We observe from the figure that wall temperature gradient increases with the increase of values of magnetic parameter (Mn).

Fig.4 (c) is presented to show the variation of wall temperature gradient $\theta'(0)$ versus visco-elastic parameter (k_1) for different values of Mn, Pr and β with different combinations values of p , Q^* and S. It is noticed from the figure that wall temperature gradient increases with increase in values of visco-elastic parameter (k_1).

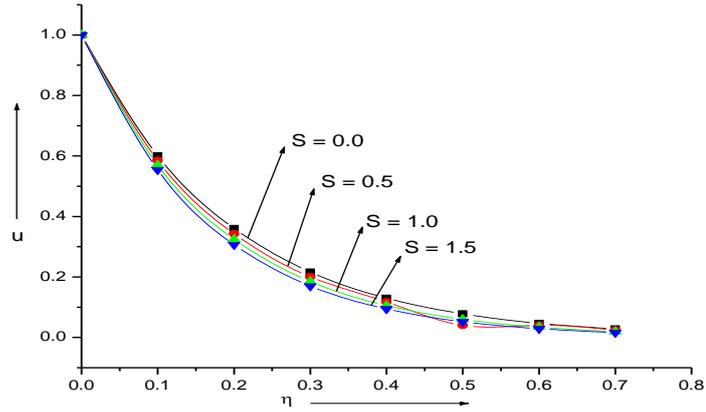


Figure1(a). Effects of S on longitudinal velocity.

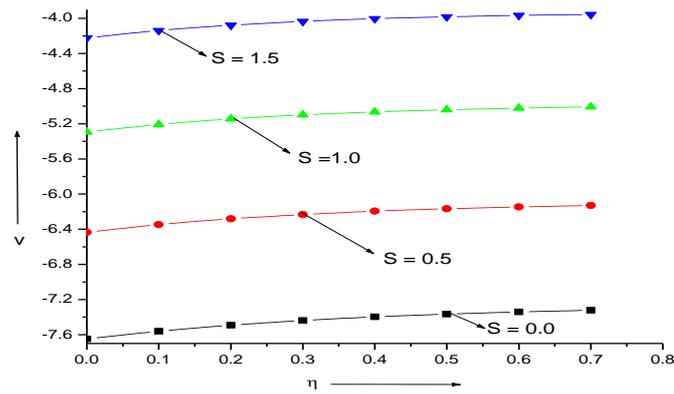


Figure 1(b). Effects of S on transverse velocity.

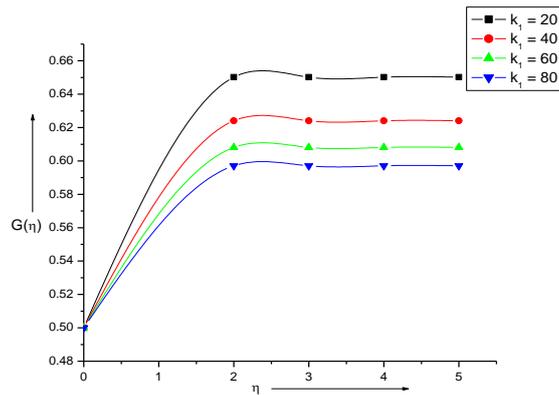


Fig.2 (a). Variation of $G(\eta)$ Vs η for $Mn=20, S=0.5$ with different values of k_1

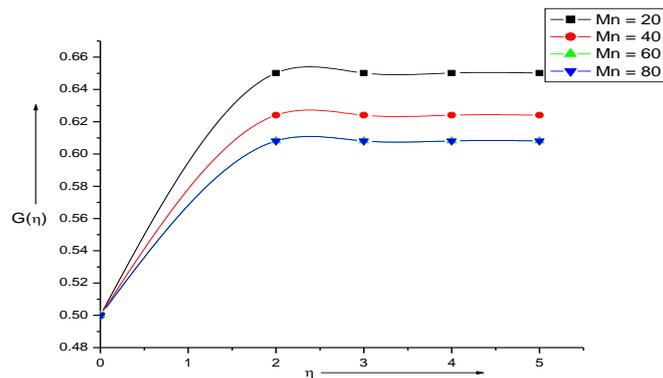


Fig.2(b). Variation of $G(\eta)$ Vs η for $k_2=20, S=0.5$ with different values of Mn

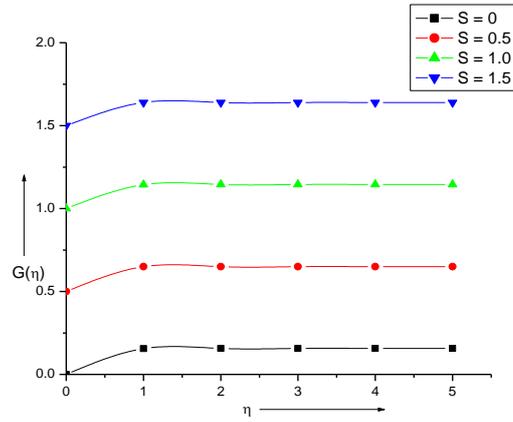


Figure 2(c). Variation of $G(\eta)$ versus η for different values of suction parameter-S

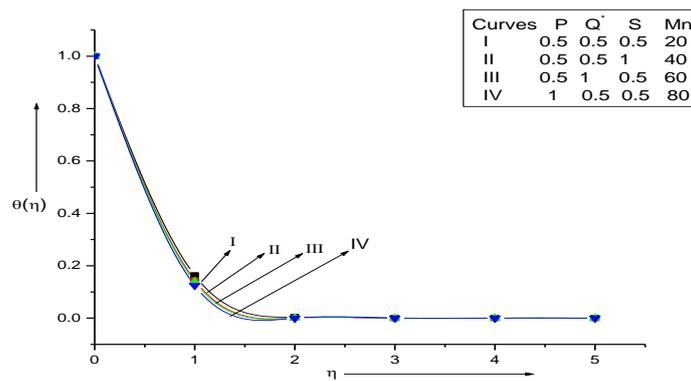


Fig. 3(a). Variation of $\theta(\eta)$ Vs η for $k_2=20, Pr=0.7, \beta=2$ with different values of Mn

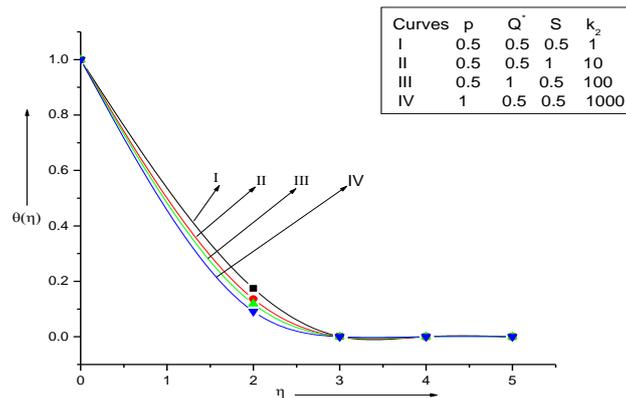


Fig.3(b). Variation of $\theta(\eta)$ Vs η for $Mn=20, Pr=0.7, \beta=2$ with different values of k_2

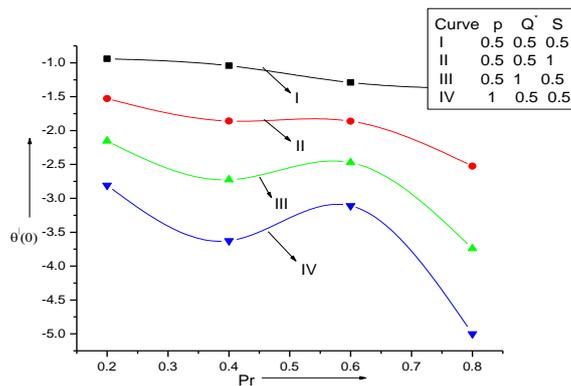


Fig.4(a). Variation of $\theta'(0)$ Vs Pr for $k_2=20, Mn=20, \beta=2$ with different values of P, Q', S

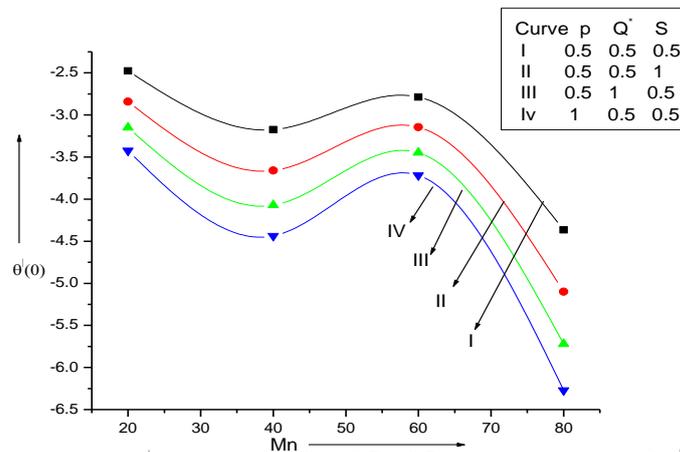


Fig.4(b) Variation of $\theta(0)$ vs Mn for $k_2=20, Pr=0.7, \beta=0.7$ and different values of P, Q, S

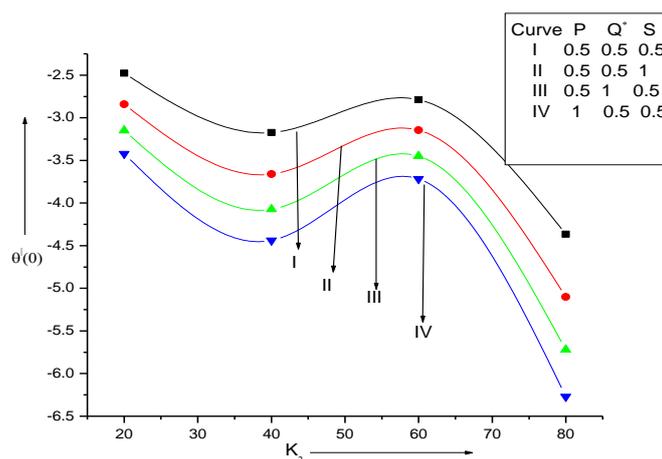


Fig.4(c).Variation of $\theta(0)$ Vs K_2 for $Mn=20, Pr=0.7, \beta=2$

REFERENCES

- Sakiadis B.C.; Boundary layer behaviour on a continuous solid surface.
- (i) Boundary layer for two dimensional and Axis-symmetric flow, IICH, E.J. 7 (1961) 26-34. (ii) Boundary layer on a continuous cylindrical surface, IICH, E-J-7 (1961) 221-232. (iii) Boundary layer on a continuous cylindrical surface. IICH E J-7 (1961) 267-274.
- Mc-Carmack P.D. and Crane L.J.; Physical fluid dynamics. Academic Press, New-York, 1973.
- Crane L.J.: Flow past a stretching plate. z angew.Math. Phys. 21 (1970) 645-647.
- Gupta P.S. and Gupta A.S.; Heat and Mass transfer on a stretching sheet with suction blowing. Canad. J. Chem. Engg. 55 (1977) 744.
- Grubka C.J. and Bobba K.M.:Heat transfer characteristics of a continuous stretching surface with variable temperature. ASME. J. Heat transfer 107 (1985) 248-250.
- Bujurke N.M., Biradar S.N., Hiremath P.S.: Second order fluid flow past a stretching sheet with heat transfer. Z. Angew.Math.Phys.38(1987) 653-657.
- Datta B.K., Roy P. and Gupta A.S.; Temperature field in a flow over a stretching that within uniform heat flux. Int. Comm. Heat transfer.12 (1985) 89.
- Chen C.K. and Char M.I.; Heat transfer on a continuous stretching surface with suction, or blowing. J. Math. Anal. Appl. 135, (1988) 568.
- Abel M.S. and Veena P.; Visco-elastic fluid flow and heat transfer in a porous medium over a stretching sheet (1998). Int. J. non-linear Mechanics Vol. 33, 531-540.
- Vajravelu K.and Rollins D.; Heat transfer in a visco-elastic fluid over a stretching sheet. J. Math. Anal. Appl. 158 (1991) 241-255.

11. Vajravelu K. and Nayeth J.: Convective heat transfer at stretching sheet. ACTA Mech.96(1-4) (1993) 47-54.
12. Anjalidevi S.P. and Thiyarajan M.; Non-linear hydro-magnetic flow and heat transfer over a surface stretching with power law velocity.
J. Heat and Mass transfer 38, (2002), 407-413.
13. Rajagopal K.R., Na T.Y. and Gupta A.S.; Flow of a visco-elastic fluid over a stretching sheet. Rheol, ACTA, 23 (1984), 213-221.
14. Sanyal D.C. and Das gupta S.; Stead flow and heat transfer of a conducting fluid caused by stretching porous wall.
Indian Journal Theo. Physics, 51 (2003) 47-58.
15. Sujitkumar Khan and Sanjayanad E.: Visco-elastic boundary layer flow and heat transfer over an exponential stretching sheet.
Int. J. of Heat and Mass transfer.48(2005) 1534-1542.
16. Soundalgekar V.M., Takhar H.S., Das V.N., Deka R.R. and Sarmah A.; Effect of variable viscosity on a boundary layer flow along a continuous moving plate with variable surface temperature.
J. Heat and Mass transfer, 40 (2004) 421-424.
17. Mahapatra T.R. and Gupta A.S.; Heat transfer in a stagnation point flow towards stretching sheet.
J. Heat and Mass transfer, 38 (2002) 517-521.
18. Bhargava R Kumar and Takhar H.S.; Mixed convection from a continuous surface in a parallel moving stream of micro-polar fluid.
J. Heat and Mass transfer, 39 (2003), 407-413.
19. Idress M.K. and Abel M.S.; Visco-elastic flow past a stretching sheet in a porous medium and heat transfer with internal heat source.
Indian journal Theo, Physics, 44 (1996) 233-244.
20. Takhar H.S. and Soundalgekar V.M.; Flow and heat transfer of micro-polar fluid past of a continuous moving porous plate.
International journal of engineering science, 23 (1985) 201-209.
21. Khan S.K., Abel M.S. and Santh R.M.; Visco-elastic MHD flow heat and mass transfer over a stretching sheet with dissipation of energy and stress work.
International journal of Heat and Mass transfer, 40 (2004) 47-57.
22. Takhar H.S, Pop I. and Soundalgekar V.M.; Dispersion of a solvable matter in porous medium channel with homogeneous and heterogeneous chemical reaction.
Revue Romanie Mechanique Appliquee, 28 (1983) 127-132.
23. Dandapat B.S and Gupta A.S.; Flow and heat transfer in visco-elastic fluid over a stretching sheet.
Int. J. non-linear. Mechanics. 24 (3) (1989), 215-219.

Magneto hydro dynamic Non-Similar solutions of a Visco-elastic fluid with Slip flow and Heat Transfer over a Non-linearly Stretching Porous Sheet

K. Ashok Kumar, P. H. Veena, V. K. Pravin,

Dept. of Mathematics, Rural Engineering College, Bhalki, Karnataka, India

Dept. of Mathematics, Smt. V.G.College for Women, Gulbarga, Karnataka, India

Dept. of Mech. Engg., P.D.A. College of Engg., Gulbarga, Karnataka, India

Abstract: The study of boundary layer flow and heat transfer of an electrically conducting visco-elastic second grade fluid in a porous medium past a stretching sheet is conducting with power law surface temperature or power law surface heat flux. The flow in boundary layer is considered to be generated solely by the linear stretching of the boundary sheet adjacent to porous medium in a magnetic field with boundary wall slip condition. The governing partial differential equations are converted into non linear ordinary differential equations by similarity transformations. The solutions of dimensionless surface temperature as well as non-similar flow and heat transfer characteristics with the governing dimensionless parameters of the problem which include a non-linear stretching sheet, viscous dissipation, internal heat generation/absorption and temperature gradient dependent heat sink, power-law index of wall temperature parameters are obtained in terms of Confluent Hypergeometric Functions (CHF) and tabulated.

The skin friction at the wall is also derived. It is observed that the suction (S), Slip parameter (L), the permeability of the medium (k₂) and the magnetic parameter (M), visco-elastic parameter (k₁) depress the longitudinal velocity magnitudes but influences positively the transverse velocity while the suction, wall temperature parameter, temperature gradient dependent heat sink (Q), lowers temperature and heat transfer distribution aiding in controlling momentum and heat transfer during material processing.

Key Words: MHD flow, stretching sheet, slip parameter, non-similar solutions, visco-elastic parameter

I. Introduction

Study of heat transfer and visco-elastic flow induced by heated stretching surfaces is often encountered in many engineering applications, such as materials manufactured by extrusion process, polymer processing, wire and fiber coating, cooling of metallic sheets or electronic chips, crystal growing.

It is well known that the flow in a boundary layer separates in the regions of adverse pressure gradient and the concurrence of separation has several undesirable effects in so far as it leads to increase in the drag on the body immersed in the flow and adversely affects the heat transfer from the surface of the body.

In context to the well-known Blasius[1] flow problem (Cortell 2005) which involves the boundary layer flow passing through a stationary flat plate. (Sakiadis, 1961) considered the boundary layer flow on a moving flat plate in quiescent ambient fluid. The afore mentioned problems are two special cases of more general studies(I-Shak et al 2007, Cortell 2007) in which flow and heat transfer of a moving sheet in the presence of a co-flowing fluid were analysed.

Very recently (Sadeghy et al 2005) studied the boundary layer of an upper convected Maxwell fluid, and the role played by the fluids elasticity on flow characteristics were analysed. Above mentioned works were with respect to the linear stretching sheet and it may be noted the stretching of the sheet may not necessarily be linear on view of this, the flow influenced by a non-linearly stretching sheet was investigated by Vajravelu (2001) and power-law or exponentially stretching sheet was studied by Ali(1995) and Elbashesly (2001) respectively. Further momentum, heat and mass transfer over an exponentially stretching surface were considered by Sanjayanand and Khan(2006). They also enclosed the effects of viscous dissipation and work done by deformation in the energy equation.

Now days the stretching sheet fluid flow is also one of the important flow fields in real world. Therefore the problem of uniqueness of a visco-elastic fluid flow over a stretching sheet has been discussed by Troy et al[1987] and Chang[1989]. Some of these visco-elastic fluids are termed as second grade fluids.. The visco-elastic property of these fluids has found in some dilute polymer solutions or in polymer fluids as mentioned by Markovitz and Colealan (1964).

For an incompressible homogeneous second grade fluid, the constitutive equation based on the postulate of fading memory suggested by Rivlin-Erickson (1955) is expressed as

$$T = -pI + \mu A_1 + \alpha_1 A_2 + \alpha_2 A_1^2 \dots\dots\dots(1)$$

Where T is the stress tensor, p is the indeterminate considered by the incompressibility pressure, μ is the dynamic viscosity, α₁ and α₂ are first and second normal stress co-efficient that are related to the materials modulus. The second grade fluid is compatible with thermo dynamics if the Helmholtz free energy of the fluid is a minimum when it is locally at rest, and further if the second grade fluid is to satisfy the classius-Dehum in equality then the co-efficient μ, α₁, α₂ must satisfy the following requirements:

$$\mu \geq 0, \quad \alpha_1 > 0, \quad \alpha_1 + \alpha_2 = 0 \dots\dots\dots(2)$$

The kinematics tensors A_1 and A_2 are defined as

$$A_1 = \nabla V + (\nabla V)^T$$

$$A_2 = \frac{dA_1}{dt} + A_1(\nabla V) + (\nabla V)^T A_1 \quad \dots\dots\dots(3)$$

where V is the velocity and $\frac{d}{dt}$ is material time derivative

Recently, the interest in that transfer for non-Newtonian fluid flows through MHD and porous medium has grown considerably due to their industrial applications such as in petroleum extrusion, enhanced oil recovery, filtration processes, in nuclear reactors packed bed reactions and many others. In this point of view Chauhan and Takhar(2002) investigated non-Newtonian coupled flow in a channel bounded by a highly porous layer.

In all the above analyse, the common feature is the assumption that the flow field satisfies the conventional no-slip condition at the stretching sheet, this assumption of no-slip is not valid and must be replaced by partial slip boundary condition, following Navier(1827) and Gad-el-Hak(1999)

$$u_i = L \frac{\partial u_i}{\partial \eta} \quad \dots\dots\dots(4)$$

Where $L = \frac{2-F}{F} \lambda$, u_i is the tangential velocity, η is the normal direction to the wall, F is the momentum accommodation co-efficient, λ is the mean free path and L is the slip length. Most of the researchers investigated about the studies with slip condition of Newtonian fluid flows and very few have worked on non-Newtonian flows with slip condition. On real aspects non-newtonian fluids such as polymer melts which often exhibit boundary wall slip and such fluids are very important from technological point of view; for example, polymer processing, artificial heart valves polishing. Labropulu et al(2004), Hayat et al(2007) and Ajadi et al(2009) discussed about the non-Newtonian fluid flows with the effects of slip condition. Ariel et al(2006) examined non-Newtonian fluid flow past a stretching sheet with partial slip. Khan et al (2008) investigated effects of slip parameters on shearing non-Newtonian fluid MHD flow through porous medium and obtained numerical solution for these typical shearing flows i.e Couette flow and generalized Couette flow with non linear slip boundary conditions.

In all the above studies the effect of temperature gradient dependent heat sink/source parameter and non-similar term parameters on heat transfer flow have not been which specially find applications in material processing industries.

In the present century, a century of technological advancement, exploration of industries using latest technologies in extrusions in manufacturing processes and melt spinning processes is taking place. In these industries the extrudate is stretched into a filament when it is drawn from the die and solidifies in the desired shape through a controlled cooling system coupled flow in a channel bounded by a stretching sheet and a highly porous medium. Garg and Rajgopal (1991), Singh et al (2006) examined the non-newtonian fluid past a wedge.

Therefore many authors including Veena et al (2010), Rafia et al (2007) have analysed problems on boundary layer flow caused by a stretching sheet with temperature gradient heat sink effects for different flow models. Similarly Veena et al (2006), Shahjahan et al(2007), Pravin et al (2006) have investigated the non similar solutions on boundary layer flow past a stretching sheet with different heat transfer parameters. Rajgopal and Gupta (1984) and Garg and Rajgopal (1991) discussed that an additional boundary condition is required since the order of momentum equation of a second grade fluid is one order

higher than that for a Navier-Stokes fluid and $\frac{\partial u}{\partial y} = 0$ as $y \rightarrow \infty$ may be taken as the augmented condition for the flow in an unbounded domain.

Thus in the present paper we are concerned not only with the natural convection over a stretching sheet but also the non-similar solutions of heat transfer flow of second grade fluid in the presence magnetic field and permeability parameter with temperature gradient dependent heat sink effects with power-law surface temperature (PST) or power-law surface heat flux (PHF). Slip flow boundary condition has been applied at the stretching sheet. In the heat equation viscous dissipation, internal heat generation or absorption, are also considered. Effects of slip parameter, permeability parameter (k_2), magnetic parameter (Mn), visco-elastic parameter (k_1), suction parameter (S), wall temperature parameter (T) temperature gradient dependent heat sink parameter (Q). Longitudinal and transverse velocity distribution for both PST and PHF cases are investigated and the results obtained are depicted graphically.

II. FORMULATION OF THE PROBLEM

In Cartesian co-ordinate system (x, y) consider two – dimensional non-similar solutions of free convection steady laminar boundary layer flow of an incompressible visco-elastic fluid caused by moving porous sheet embedded in a porous medium in presence of a temperature gradient dependent heat sink. The porous sheet is subjected to a constant suction velocity normal to the wall and uniform magnetic field. The x-axis is taken along the wall in the direction of motion of the flow and y-axis perpendicular to it. Let the components of velocity be u and v along x and y directions respectively. It is envisaged that the sheet issues from a thin slit at the origin (0, 0) and the speed at a point on the plate is proportional to its distance from the plate but the boundary layer approximations holds true.

Under the above mentioned assumptions and following Vafai and Tien`s [28] model, the steady state boundary layer equations of mass momentum and energy are given by

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad \dots (5)$$

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \nu \frac{\partial^2 u}{\partial y^2} + \frac{\alpha_1}{\rho} \left[\frac{\partial}{\partial x} \left(u \frac{\partial^2 u}{\partial y^2} \right) + \frac{\partial u}{\partial x} \cdot \frac{\partial^2 v}{\partial y^2} + v \frac{\partial^3 u}{\partial y^3} \right] - \frac{\nu}{K'} u - \frac{\sigma B_0^2 u}{\rho} \quad \dots(6)$$

and

$$\rho C_p \left(u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} \right) = k \frac{\partial^2 T}{\partial y^2} + Q(T - T_\infty) + Q' \frac{\partial T}{\partial y} + \alpha_1 \frac{\partial u}{\partial y} \cdot \frac{\partial}{\partial y} \left(u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} \right) + \mu \left(\frac{\partial u}{\partial y} \right)^2 \quad \dots(7)$$

The supplementary terms in the momentum equations namely the Darcian body force term

$$-\frac{\nu}{K'} u \text{ and magnetic conductivity term } -\frac{\sigma B_0^2 u}{\rho}$$

are linear in terms of the x-direction velocity u i.e they are parallel to the direction of the stretching motion. T-the temperature, ρ - the density , $\nu = \frac{\mu}{\rho}$ - the kinematics' viscosity, k_2 the permeability parameter, Mn be the magnetic parameter , α_1 , the non-Newtonian parameter, C_p , the specific heat , Q' - the uniform specific heat source or sink, k- the thermal conductivity and T_∞ - the temperature at infinity.

The supplementary term in the energy equation namely the temperature gradient dependent heat sink $-Q' \frac{\partial T}{\partial y}$ and $Q'(= -CQG)$ which is the volumetric rate is a linear function of the temperature field.

It is assumed that the contribution due to the normal stress is of the same order of magnitude as that due to the shear stress since the flow is driven solely by stretching the sheet and the pressure gradient is assumed to be absent.

The appropriate boundary conditions for the momentum problem are

$$u(x, y) - bx = L \frac{\partial u}{\partial y}, \quad v = -V_0 \quad \text{at} \quad y = 0$$

$$u \rightarrow 0, \quad \frac{\partial u}{\partial y} \rightarrow 0 \quad \text{as} \quad y \rightarrow \infty \quad \dots(8)$$

where $b > 0$ is a constant.

III. SOLUTION OF MOMENTUM PROBLEM

To solve equation (6), we postulate a solution by introducing the following similarity transformations:

$$u = bx G'(\eta), \quad v = -\sqrt{bv} G(\eta) \quad \text{and} \quad \eta = \sqrt{\frac{b}{\nu}} y \quad \dots(9)$$

Obviously with this choice of velocity variables u and v , equation (1) of continuity is identically satisfied.

Substituting (9) in equation (6) we obtain

$$G'^2(\eta) - GG'' = G''' + [2GG''' - G''^2 - GG^{IV}] - (K_2 + Mn)G' = 0 \quad \dots(10)$$

where $k_2 = \frac{\nu}{K'b}$ permeability parameter, $Mn = \frac{\sigma B_0^2}{\rho b}$ - Magnetic parameter,

$$k_1 = \frac{\alpha_1 b}{\mu} = \text{visco-elastic parameter}$$

By use of (9), the boundary conditions (8) corresponding to equation (6) reduce to:

$$G(\eta) = S, \quad G'(\eta) = 1 + \gamma G''(\eta) \quad \text{at } \eta = 0 \quad \dots(11)$$

$$G'(\eta) \rightarrow 0, \quad G''(\eta) \rightarrow 0 \quad \text{as } \eta \rightarrow \infty$$

where $\gamma = L \left(\frac{b}{\nu} \right)^{1/2}$ is the slip parameter

and $S = \frac{V_0}{\sqrt{\nu b}}$ is the suction parameter.

The solution of equation (10) subjected to the boundary conditions (11) is

$$G(\eta) = A_1 + B_1 e^{-\alpha \eta} \quad \dots(12)$$

And hence $G'(\eta) = -\alpha B_1 e^{-\alpha \eta}$

where A_1 and B_1 are constants to be determined such that it satisfied (10) under the boundary conditions (11). Thus

$$G(\eta) = \frac{1}{\alpha(1+\alpha\gamma)} (1 - e^{-\alpha\eta}) \quad \dots(13)$$

and here $\alpha \neq \frac{1}{\gamma}$

It satisfies all the boundary conditions (11) and it is an exact solution provided one root α is real of the following cubic equation

$$\gamma \alpha^3 + (1 + k_1) \alpha^2 - (\gamma(k_2 + Mn)) \alpha - (1 + k_2 + Mn) = 0 \quad \dots(14)$$

SKIN FRICTION

The co-efficient of skin friction at the stretching sheet ($\eta=0$) is obtained as

$$\tau = \frac{\tau_{yx}(0)}{\rho u_w^2 / 2} = \frac{-2\alpha [1 + \alpha\gamma + 3k_1]}{Re_x^{1/2}} \quad \dots(15)$$

where $Re_x = \frac{bx^2}{\nu}$

IV. SOLUTION OF HEAT TRANSFER PROBLEM

The boundary conditions for temperature field depend on the type of heating process under consideration.

The prescribed surface temperature (PST) case:

In this case the boundary conditions are

$$T = T_w = T_\infty + A \left(\frac{x}{l} \right)^2 \quad \text{at } y = 0$$

$$T \rightarrow T_\infty \quad \text{as } y \rightarrow \infty \quad \dots(16)$$

where A is a constant and $T_w = T_\infty + A\left(\frac{x}{l}\right)^2$ is the wall temperature function defining non-isothermal behavior in terms of quadratic power law. At the leading edge of the boundary layer $x = 0$ the wall temperature reduces to an isothermal law i.e. $T_w \rightarrow T_\infty$.

To solve heat equation (7), we introduce the following non-dimensional variable θ :

$$\theta = \frac{T - T_\infty}{T_w - T_\infty} \quad \dots(17)$$

Further introducing (16) and (17), equation (7) reduces to

$$\theta'' + \text{Pr} (1 + Q^*) G \theta' - \text{Pr} (2G' - \beta) \theta = -Ec$$

$$\text{Pr} \left[-k_1 G'' (G' G'' - G G''') + G''^2 \right] \quad \dots(18)$$

where $\text{Pr} = \frac{\mu C_p}{k}$ - the Prandtl number ; $\beta = \frac{Q}{b \rho C_p}$ - viscous dissipation

$Ec = \frac{b^2 l^2}{A C_p} \sqrt{b}$ - Eckert number ; $Q^* = Q \sqrt{\frac{b}{\nu}}$ - Internal heat generation

and the boundary conditions (16) transform to

$$\theta(\eta) = 1 \quad \text{at} \quad \eta = 0$$

$$\theta(\eta) \rightarrow 0 \quad \text{as} \quad \eta \rightarrow \infty \quad \dots(19)$$

To obtain the solution of equation (18), we introduce a change of variable ξ defined as

$$\xi^{-1} = \frac{-\alpha^2 e^{\alpha \eta}}{\text{Pr}(1+Q)} \quad \dots(20)$$

Hence with the help of (20) equation (18) transforms to

$$\xi \theta'' + [1 - a_0 - \xi] \theta' + \left[S_1 + \frac{\text{Pr}}{\alpha^2} \beta \xi^{-1} \right] \theta = \frac{-Ec}{\text{Pr}} \left[1 - k_1 \left\{ 1 - \frac{v_0}{\sqrt{b\nu}} \alpha \right\} \right] \frac{\alpha^4}{(1+Q^*)^2} \xi \quad \dots(21)$$

where $a_0 = \frac{\text{Pr}(1+Q^*)}{\alpha^2} \left\{ 1 - \frac{v_0}{\sqrt{b\nu}} \alpha \right\}$, $S_1 = \frac{2}{1+Q^*}$ and the corresponding boundary conditions are

$$\theta \left(\xi = \frac{\text{Pr}(1+Q)}{\alpha^2} \right) = 1, \quad \theta(\xi=0) = 0 \quad \dots (22)$$

Equation (21) is in standard confluent hypergeometric equation form that is Kummer's equation [see Sanyal and DasGupta[2003] and Abramowitz and Stegun (1965) and the solution of equation (21) with respect to boundary conditions (22) is obtained as

$$\theta(\xi) = C_1 \xi^k M^\Gamma [K - S_1, 1 + b_0, \xi] - \frac{Ec \left[1 - k_1 \left(1 - \frac{V_0}{\sqrt{b\nu}} \alpha \right) \right] \alpha^4}{\text{Pr} \left[4 - 2a_0 + \frac{\text{Pr}}{\alpha^2} \beta \right]} \xi^2 \quad \dots(23)$$

where $K = \frac{a_0 + b_0}{2}$, $a_0 = \frac{\text{Pr}(1 + Q^*)}{\alpha^2} \left\{ 1 - \frac{v_0}{\sqrt{b\nu}} \alpha \right\}$, $b_0 = \sqrt{a_0^2 - \frac{4\text{Pr}\beta}{\alpha^2}}$

Solution (23) in terms of the similarity variable η is expressed as

$$\theta(\eta) = \xi e^{-\alpha k \eta} \frac{M \left[K - S_1, 1 + b_0, \frac{-\text{Pr}(1 + Q^*)}{\alpha^2} e^{-\alpha \eta} \right]}{\left[K - S_1, 1 + b_0, \frac{-\text{Pr}(1 + Q^*)}{\alpha^2} \right]} + B_1 e^{-2\alpha \eta} \quad \dots(24)$$

where $\xi = 1 - B_1$, $B_1 = \frac{-Ec\alpha^2 \left[1 - k_1 \left(1 - \frac{v_0 \alpha}{\sqrt{c\nu}} \right) \right] \text{Pr}(1 + Q^*)^2}{[4\alpha^2 - 2a_0 \alpha^2 + \beta \text{Pr}]}$

The dimensionless temperature gradient $\theta'(0)$ derived from equation (24) is as follows

$$\theta'(0) = A_2 \left[k_3 \alpha \left(\frac{K - S_1}{1 + b_0} \right) M [K - S_1 + 1, b_0 + 2, -k_3] - K \alpha M [K - S_1, 1 + b_0, -k_3] \right] - 2B_1 \alpha \quad \dots(25)$$

where $k_3 = \frac{\text{Pr}(1 + Q^*)}{\alpha^2}$, $A_2 = \frac{A_1}{M [K - S_1, 1 + b_0, -k_3]}$

The dimensionless rate of heat transfer at the stretching sheet ($\eta = 0$), characterized by the Nusselt number is given by

$$Nu = \frac{-k \left(\frac{\partial T}{\partial y} \right)_{y=0}}{k(T_w - T_\infty)} x = \text{Re}_x^{1/2} \theta'(0) \quad \dots(26)$$

The prescribed power law surface heat flux (PHF)

In this case we take the following boundary conditions.

$$-K \frac{\partial T}{\partial y} = q_w = \frac{B}{K} \left(\frac{x}{l} \right)^2 \quad \text{at } y = 0 \quad \dots(27)$$

and $T \rightarrow T_\infty$ as $y \rightarrow \infty$

we define $T - T_\infty = \frac{B}{K} \left(\frac{x}{l} \right)^2 \left(\frac{\nu}{c} \right)^{1/2} g(\eta)$... (28)

on substituting (28) and (17) in to equation (7), we obtain

$$g'' + \text{Pr} (1 + Q^*) G g' - \text{Pr} (2G' - \beta) g = -Ec \text{Pr} \left[-k_1 G'' (G' G'' - GG''' + G''^2) \right] \quad \dots(29)$$

where $Ec = \frac{K b^2 l^2}{ACp} \sqrt{\frac{b}{\nu}}$, $k_1 = \frac{k_0 b}{\nu}$ - and all other parameters are the same as before

Using the transformation

$$\xi = \frac{-\text{Pr}(1+Q^*)}{\alpha^2 e^{a\eta}} \quad \dots(30)$$

Equation (29) takes the form which reduces to the following confluent hypergeometric equation

$$\xi g'' + [1 - a_0 - \xi] g' + \left[S_1 + \frac{\text{Pr}}{\alpha^2} \beta \xi^{-1} \right] g = \frac{-Ec}{\text{Pr}} \left[1 - k_1 \left\{ 1 - \frac{\nu_0}{\sqrt{c\nu}} \alpha \right\} \right] \frac{\alpha^4}{(1+Q^*)^2} \xi \quad \dots(31)$$

If we take

$$g(\xi) = g_c(\xi) + g_p(\xi) \quad \dots(32)$$

The corresponding boundary conditions (27) reduce to

$$g'(0) = -1 \quad \text{and} \quad g(0) = 0 \quad \dots(33)$$

Solving equation (31) under the boundary conditions (33) and using (32), we obtain the solution of (31) in terms of η as

$$g(\eta) = B_1 e^{-2\alpha\eta} + B_2 e^{-\alpha K \eta} M \left[K - S_1, 1 + b_0, -k_3 e^{-\alpha\eta} \right] \quad \dots(34)$$

where

$$B_2 = \frac{(-1 + 2\alpha B_1)}{\{-\alpha K M [K - S_1, 1 + b_0, -k_3] + B_3 M [K - S_1 + 1, 2 + b_0, -k_3]\}}$$

$$B_3 = \frac{\text{Pr}(1+Q^*)}{\alpha} \left(\frac{K - S_1}{1 + b_0} \right)$$

The dimensionless wall temperature is derived as

$$T_w - T_\infty = \frac{B}{k} \left(\frac{x}{l} \right)^2 \left(\frac{\nu}{c} \right)^{1/2} g(0)$$

V. Results and Discussion

The study of boundary layer flow behavior and heat transfers of a visco-elastic fluid (Walter liquid B) is considered in the presence magnetic field and porous medium adjacent to the stretching sheet with two different types of heating processes namely power-law surface temperature (PST) or power-law temperature gradient and power-law surface wall heat flux (PHF) or power-law wall temperature are considered. In addition the account of viscous dissipation, internal heat generation or absorption effect and temperature dependent gradient heat sink/source term are also considered. Several closed form solutions for the velocity and temperature fields are obtained.

Figures (1) and (2) show the variation of $G(\eta)$ and $G'(\eta)$ with the similarity variable η for different values of various parameters such as non-dimensional permeability parameter k_2 magnetic parameter Mn visco-elastic (k_1) and slip parameter (γ). It is clear from these figures that the flow velocity increases as expected. That is G and G' increase with

decrease in permeability parameter $k_2 = \left(\frac{\nu/c}{k_0} \right)$ implying increasing permeability k_0 of the porous medium causes faster

flow. Flow also increases with the increase of magnetic parameter Mn and non-Newtonian parameter k_1 however Slip parameter causes decrease in the values of both G and G' for all values of η in the boundary layer.

Fig(3) depicts $|f''(0)|$ against the slip parameter γ for various values of all the other parameters k_2 , Mn and k_1 . It is observed from the figure that the magnitude of Skin friction i.e. the magnitude of dimensionless surface velocity gradient $|f''(0)|$ decreases with slip parameter (γ) or non-Newtonian parameter k_1 which further implies that the effect of γ or k_1 is to decrease the power needed to stretch the sheet. Further the power needed to stretch the sheet is also reduced by increasing the permeability k_2 and effect of magnetic field Mn of the flow medium. The magnitude of slip, that is $(1 - G'(0))$ increases with the increase in the values of γ , because the frictional resistance between the stretching of the sheet imposes less motion of the fluid as slip parameter γ increases.

Figures (4a) and (4b) shows the variation of temperature profiles both in PST case and PHF case respectively with η for various values of the parameters. From these figures it is clear that for a given position η , the temperature decreases with an increase in the values of Prandtl number Pr with fixed values of Ec , k_1 and β in the both the cases of PST and PHF. On the other hand, the slip parameter γ has opposite effects on the temperature profiles. It may be described physically that the thermal characteristics are more influenced by the slip factor than by those other flow parameters. We see that as the value of slip parameter γ increases, the temperature distribution in boundary layer gets increased. Further both $\theta(\eta)$ and $g(\eta)$ decrease at all values of η with an increase in the values of Prandtl number Pr and thus the thermal boundary layer thickness decreases.

Fig (5a) and (5b) depict the variations in the temperature field with η for different values of Pr effects of different combinations of suction parameter S , Prandtl number Pr , temperature gradient dependent heat sink parameter Q' and with fixed values of Eckert number Ec , visco-elastic parameter k_1 . Maximum temperature corresponds to the curve- I for which $S = 0.5$, $Pr = 0.5$ and $Q = 0.5$. This physically implies that the union of weak suction with low wall temperature and weak heat sink. As expected the temperature is lowered for rise in S , Pr , Q and γ – slip parameter. However for the second curve II the S values are higher than for curve –III and curve IV. This shows that S impinges relatively less effects on decaying the temperature in comparison with the values of Pr . And Q . Such a scenario explains that a stronger heat sink plays more dominant role for lowering the temperature. These present results in general are well in agreement with the earlier studies of Sanyal and Das Gupta [2003], and Elbasheshy and Bazid [2004] in both the cases of PST and PHF.

Fig (6a) and (6b) shows the variations in temperature profiles for different values k_1 for various combinations of the parameters S , Pr , Q , γ , k_2 . From both the figures it may be concluded that the thermal characteristics are more impressed by the slip parameter. We see that for increasing values of non-Newtonian parameter k_1 and slip parameter γ , the temperature in the thermal boundary layer gets increased, resulting in a increase in the thermal boundary layer thickness. Further the heat generation due to viscous dissipation and heat source is characterized by it is also observed that there is a temperature overshoot near the stretching surface with an increase in the Eckert number, in fact it means that there is significant heat generation due to fluid friction near the sheet.

Fig (7) represents the results of the rate of heat transfer- $\theta'(0)$ and $g(0)$ Vs γ at the sheet for various values of Eckert number Ec . The effect of Eckert number Ec is to reduce the rate of heat transfer in both the case of PST in absolute sense it may be explained that for small values of Ec the rate of heat transferred is occurred from the stretching sheet to the fluid. Further it is reduced with the increase in the value of slip parameter γ . However again if the Eckert number values are allowed to be large, the rate of heat transfer decreases to its minimum value and then changes sign and finally its magnitude increases. In this scenario heat transfer takes place from the fluid region to the sheet because Ec is large enough to generate heat in the fluid is large. It is also seen that for higher values of Ec this change of sign occurs at greater value of the slip factor γ .

Figure (8) is drawn to explain the observation for the dimensionless temperature distribution $g(0)$ for fixed values of γ , β and Ec . $g(\eta)$ increases at all the values of η in the boundary layer however it reduces by the visco-elastic parameter k_1 , Prandtl number Pr or the porosity and magnetic field of the medium. Physically it is apparent that the surface temperature $g(0)$ in PHF case increase with the values of Eckert number Ec and internal heat generation β but $g(0)$ decreases with increasing values of Pr , k_1 or permeability k_2 and Mn . Further slip factor γ also causes an increase in the value of $g(0)$ as seen from the figure (8).

VI. Conclusions

The main conclusions of the study are as follows.

- i. The slip parameter γ , suction parameter S , have substantial effect on the flow and heat transfer process. The longitudinal velocity is maximum at $\eta = 0$ for all values of γ and S and decreases rapidly with increase in η and far away from the surface of the sheet.
- ii. The longitudinal velocity u increases with the decrease in the values of permeability and magnetic parameters
- iii. The transverse velocity normal to the stretching sheet tends to a constant negative value and this inflow towards the sheet from the ambient fluid decays with increasing the values of γ and S .
- iv. The transverse velocity is minimum at $\eta = 0$ for values of γ , S , k_2 and Mn and slowly start increasing with increase in values of η .
- v. The magnitude of skin friction decreases with increasing slip parameter γ , non-Newtonian parameter k_1 and increases by the permeability and magnetic parameters that is the flow in the boundary layer increases with increase in permeability of the porous medium and magnetic effect or viscid-elasticity but decreases by the slip parameter γ .

However an increase in suction parameter S or distance in the direction of the longitudinal velocity decreases the skin friction .

- vi. An increase in in slip parameter γ , suction parameter S , wall temperature parameter or heat skin parameter or temperature gradient dependent heat sink parameter Q^* results in lowering the temperature field steadily.
- vii. The effect of the slip parameter γ is to reduce the heat transfer rate. The visco-elasticity of the Walter’s liquid (B') model enhances the rate of heat transfer in PST takes place from the fluid to the stretching sheet when Ec is large enough rather than from the sheet the fluid when Ec is small.

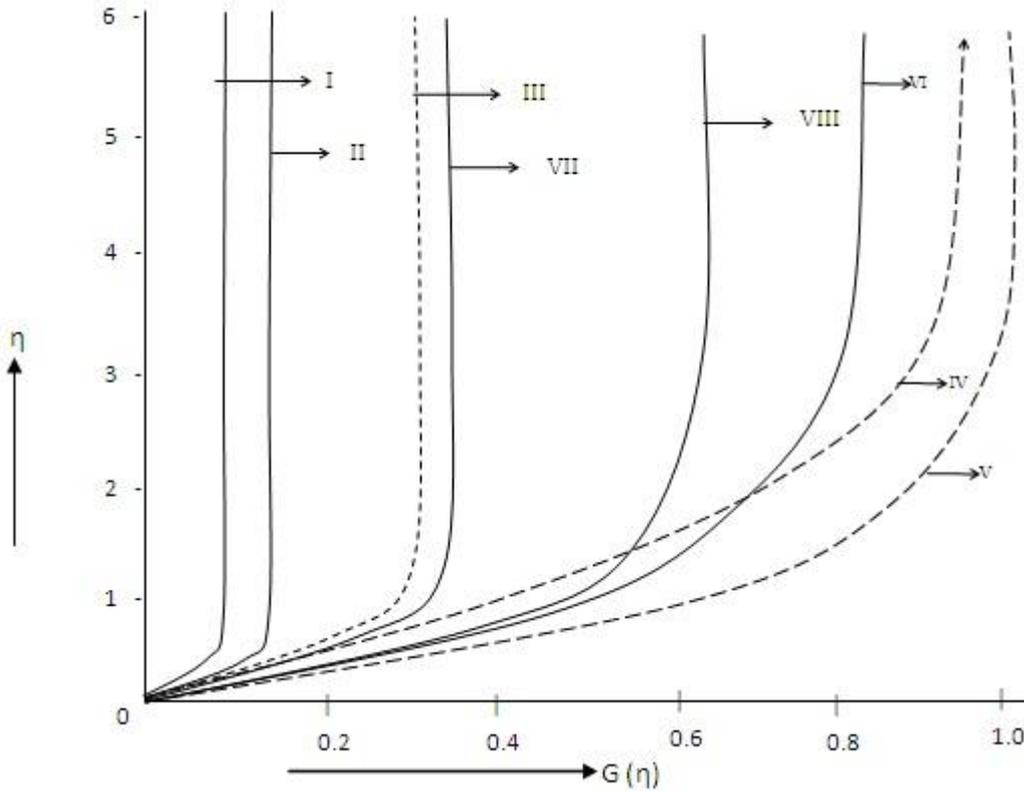
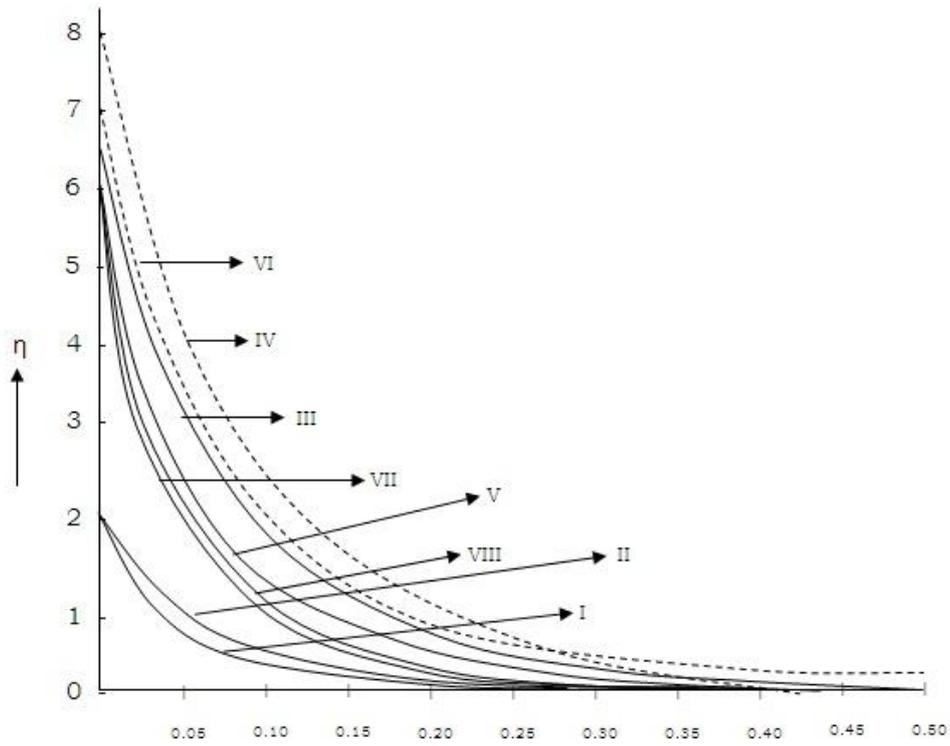


Fig 1 : Longitudinal velocity profiles $G(\eta)$ Vs η for various combinations of values of k_1, k_2, Mn and γ

Curve	k_1	k_2	Mn	γ
I	1	10	10	1
II	1	5	5	1
III	1	1	1	1
IV	1	0	0	1
V	1	1	1	0
VI	1	1	1	0.3
VII	1	1	1	0.5
VIII	0	1	1	1

Fig(2): Transverse velocity profiles $G'(\eta)$ Vs. η for various combinations of the values of k_1, k_2, Mn , and γ



Curve	k_1	K_2	Mn
I	1	10	10
II	0	1	1
III	0.1	1	1
IV	1	1	1
V	1	0	0

Fig (3) Skin friction co-efficient $|f''(0)|$ Vs slip parameter γ

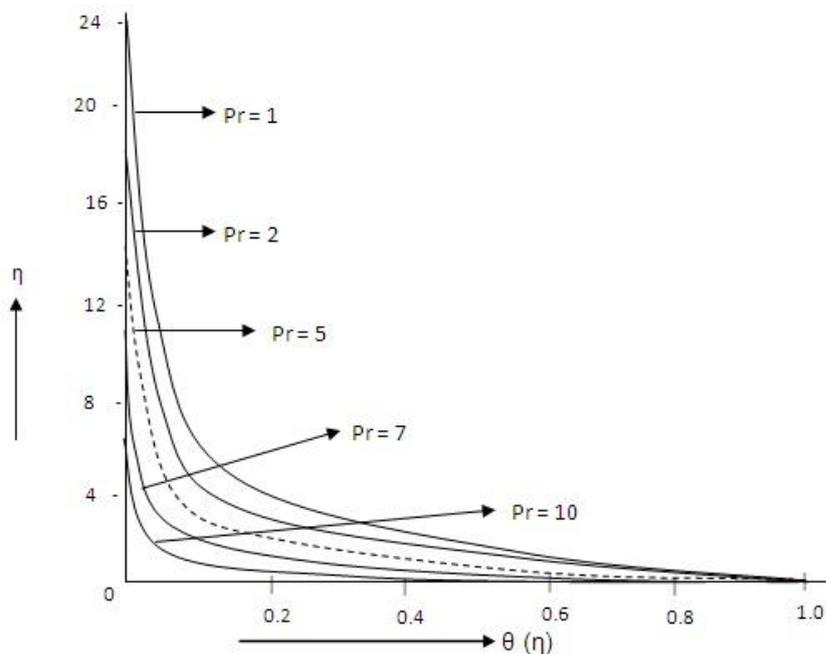


Fig (4a) Temperature profiles $\theta(\eta)$ Vs. η for different values of Prandtl number Pr and for fixed values of $k_1 = 1$, $k_2 = 10$, $Mn = 10$, $\gamma = 1$ (PST case) and $Ec = 0.1$

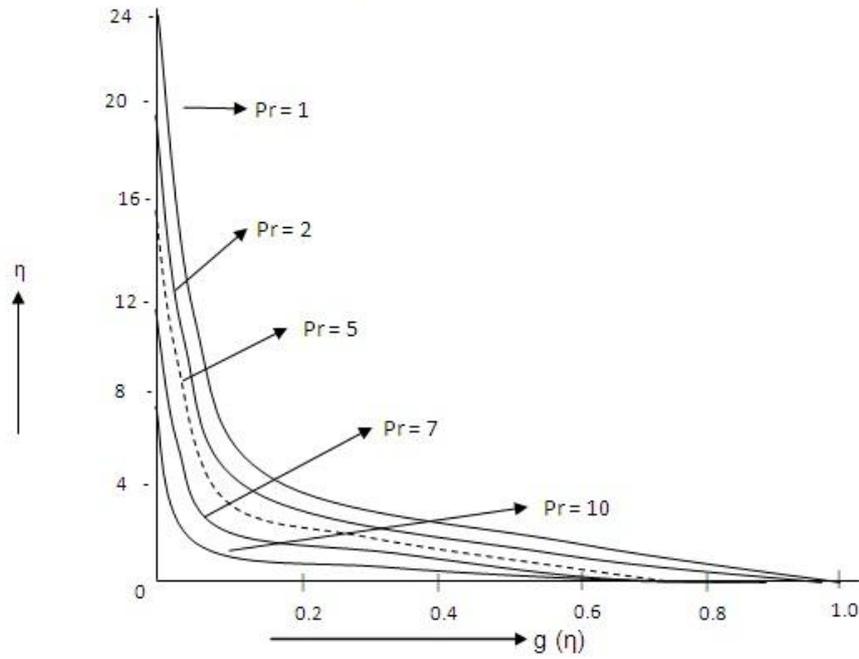


Fig (4b) Dimensionless temperature profiles $g(\eta)$ Vs. η for different values of Prandtl number Pr. (PHF-Case)

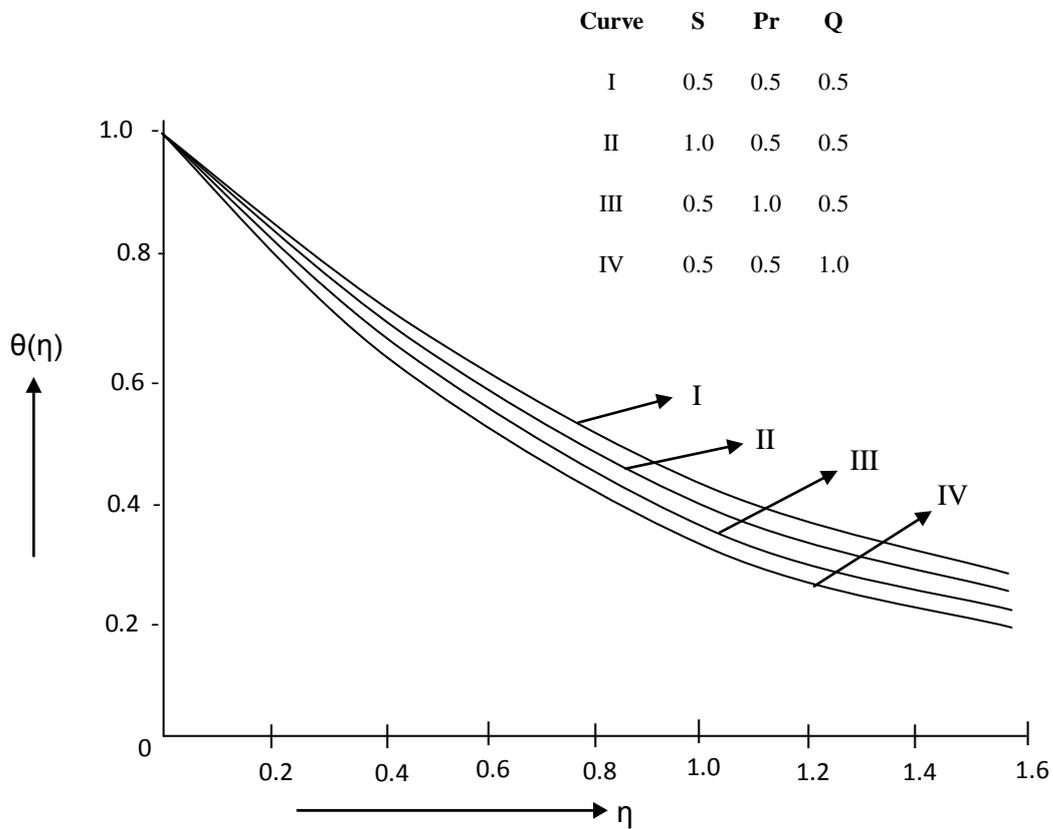
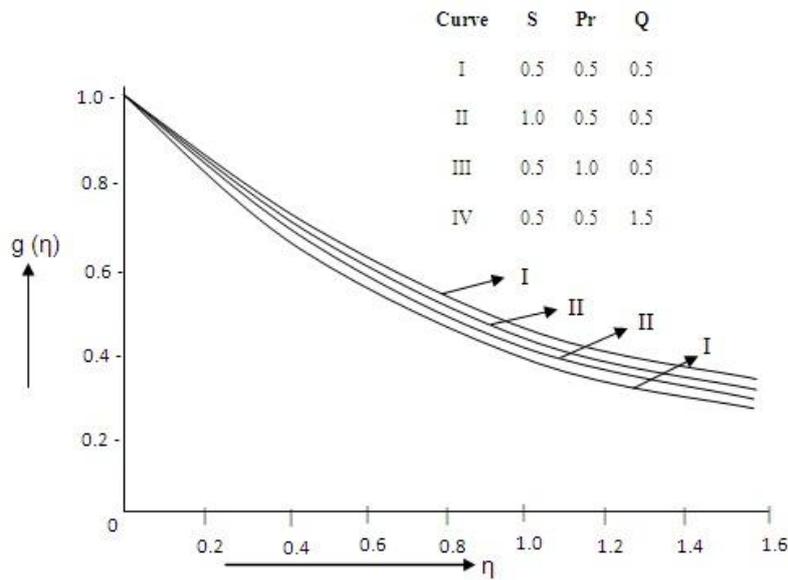


Fig (5a) Dimensionless temperature profiles $\theta(\eta)$ Vs. η for various combinations of S, Pr and Q and fixed values of $k_1 = 1$, $k_2 = 10$, $Mn = 10$, and $Ec = 0.1$ (PST - case)



Fig(5b) Dimensionless temperature profiles $g(\eta)$ Vs. η for different combinations of S, Pr and Q and fixed values of $k_1 = 1$, $k_2 = 10$, $Mn = 10$, and $Ec = 0.1$ (PHF – case)

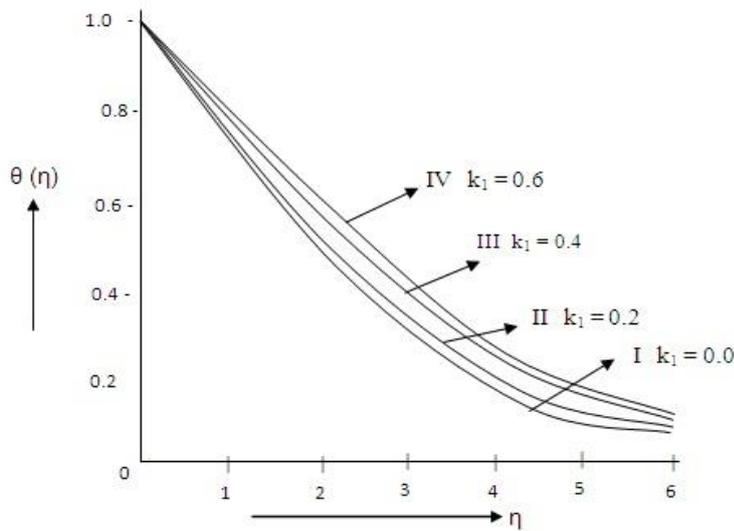
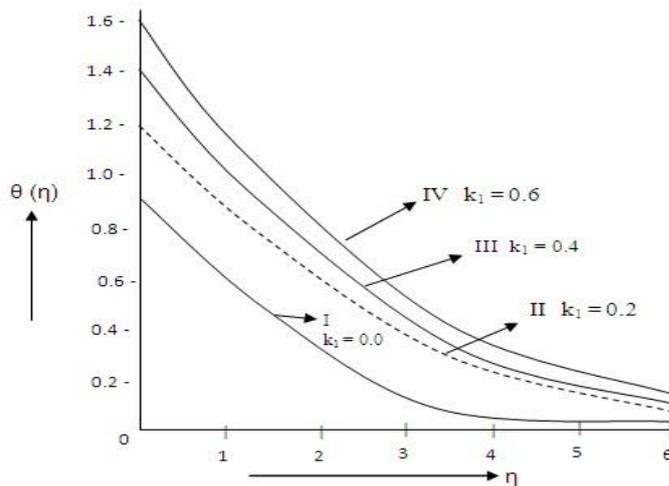


Fig (6a): Temperature profiles $\theta(\eta)$ Vs. η for various values of visco-elastic parameter k_1 with $Ec=0.5$, $Pr=1.0$, $\beta=-0.03$ in PST - case



Fig(6b): Temperature profiles $g(\eta)$ Vs. η for various values of k_1 with all other parameters as in Fig(6a). PHF - case

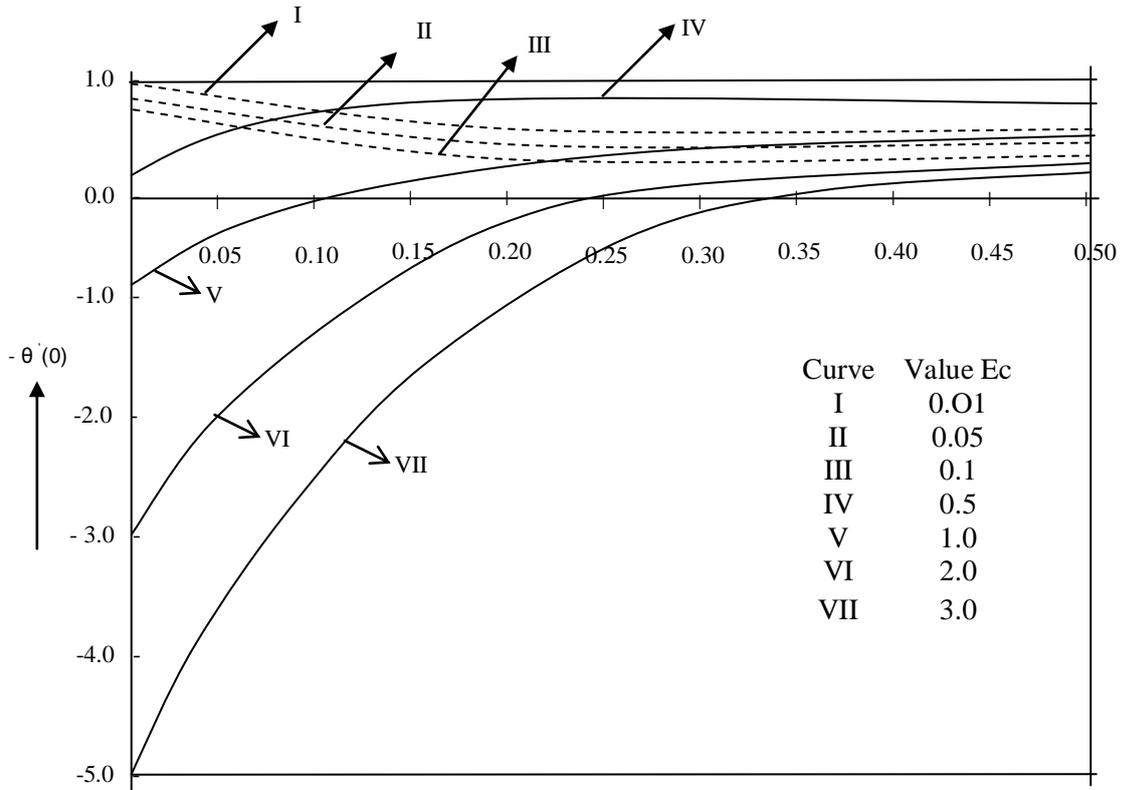
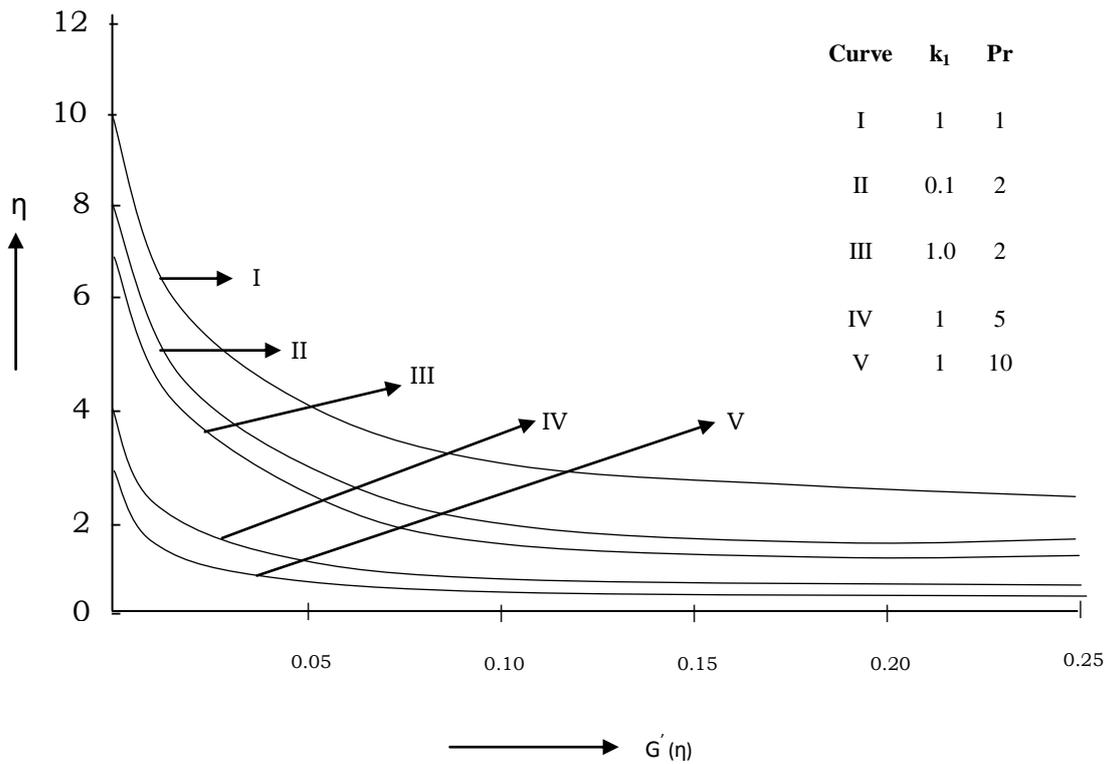


Fig (7): Dimensionless Temperature gradient $\theta'(0)$ at the stretching sheet for various values of Ec and fixed values of $k_1=1$, $k_2=10$, $Mn=10$, $Pr=1$, $\beta=-0.1$



Fig(8): Dimensionless wall temperature $g(0)$ Vs. η in PHF case for $\gamma=1$, $k_2=1$, $Mn=1$, $\beta=-0.1$ and $Ec=0.1$ for various values of combinations of k_1 and Pr

REFERENCES

1. Cortell, R (2005): Flow and heat transfer of a fluid through a porous medium over a stretching surface with internal heat generation/absorption and suction/blowing Fluid Dynamics : Res., 37, 231-245.
2. Sakiadis, B.C (1961) : Boundary layer behaviour on continuous solid surfaces: A. I. Ch. E. J., 7, 26 - 28.
3. Ishak , A. , Nazar , R., pop. : (2007) : Boundary layer on a moving wall with suction and injection. Chin. Phys. Lett. 24 (8) , 2274-2276.
4. Garg, V.K. and Rajgoal, K.R. (1991). Flow of non-Newtonian fluid past a wedge. Acta Mech., 88, 113-123.
5. Sadeghy, K., Najafi , A. H ., Saffaripour, M., (2005): Sakiadis flow of an upper- convicted Maxwell – fluid. Int. J. Non-linear Mech. 40, 1220 - 1228
6. Vajravelu , K. , (2001) : Viscous flow over a non-linearly stretching sheet. Appl. Math. Comput. 124, 281-288
7. Ali , M.E.(1995) : On thermal boundary on a porous law stretched surface with suction or injection. Int. J. Heat Mass Flow. 16, 280-290
8. Elbashbeshy, E.M.A.(2001) : Heat transfer over an exponentially stretching continuous surface with suction Arch. Mech. 53, 643-651
9. Sanjayanand, E., Khan S.K., (2006) : On heat and mass transfer in a visco-elastic boundary layer flow over an exponentially stretching sheet *Int. J. Thermal. Sciences.* 45 , 819-828
10. Troy , W. C. , Overman , E. A., Ermentrout, G.B., and Keener, J. P., (1987) :Uniqueness of flow of a second order fluid past a stretching sheet. *Quart. Appl. Math.*, 44 , 753-755
11. Chang, W. D., (1989) : The non- uniqueness of the flow of a visco-elastic fluid over a stretching sheet. *Quart. Appl. Math.* , 47 , 365 – 366.
12. Markvitz, H and Coleman, B. D. (1964): Advances in Applied Mechanics. Vol 8. Academic press, New York.
13. Rivlin, R. S and Erickson, J. L. (1955): Stress deformation relation for isotropic materials J. Rat. Mech. Anal. , 4, 323-325.
14. Chauhan , D.S. and Takhar, P.K. (2002) : Non-Newtonian stretching coupled flow in a channel bounded by a highly porous layer. Modeling Measurement and control, B , 71 (4), 33-40.
15. Navier, C.L.M.H. (1827). Sur les lois du mouvement des fluides, Mem, Acad. R.Sci. Inst. Fr., 6, 389-440.
16. Gad-el-Hak, M. (1999). The fluid mechanics of micro devices-The Free Scholar Lecture, J. Fluid Eng- T.ASME, 121, 5-33.
17. Labropulu, F., Husain, I. and Chinichain, M. (2004). Stagnation-point flow of the Walters' B'fluid with slip. IJMMS, 61, 3249-3258.
18. Hayat, T., Khan, M. and Ayub, M. (2007). The effect of the slip condition on flows of an Oldroyd 6-constant fluid. J. Comput. Appl. Math., 202, 402-413.
19. Ajadi, S.O., Adegoke, A. and Aziz, A. (2009). Slip boundary layer flow of non-Newtonian fluid over a flat plate with convective thermal boundary conditions. Internationa J. of Non-Linear Science, 8(3), 300-306.
20. Ariel, P.D., Hayat, T. and Asghar, S. (2006). The flow of an elastico- viscous fluid past a stretching sheet with slip. Acta Mech., 187, 29-35.
21. Khan, M., Hayat, T. and Wang, Y. (2008). Slip effects on shearing flows in a porous medium. Acta Mech. Sin., 24, 51-89.
22. Singh. A. K : Heat Transfer and Boundary Layer Flow past a stretching porous Wall with Temperature gradient Dependent Heat sink JEHMT, 28 (2006), 109-225.
23. Veena, P.H., Pravin. K., and J.K.Nagbhooshan : MHD Visco-elastic Flow and Heat Transfer with Temperature Gradient Dependent Heat Sink in a Porous Medium past a stretching Sheet. Int.J.Math. & Applications 3, (2010), 65-81
24. Veena,P.H., V.K.Pravin and Rafia Begum : Free and Forced MHD Convective Mass Transfer Flow with Heat Source and Thermal Diffusion embedded in a porous medium with Variable Thermal Conductivity. Int. J. Mathematics and Analysis , 5, (2007) 267-280
25. Veena,P.H., Pravin, V.K., and Shahjahan, M : Non-Similar Solutions for Heat and Mass Transfer Flow in an Electrically conducting Visco-elastic fluid over a Stretching Sheet embedded in a porous Medium. *Int.J.Modern Mathematics*, 2 (2007) 9-26.
26. Veena,P.H., V.K.Pravin., and Padashetty, S. : Non-Similar Solutions in Visco-Elastic MHD Flow and Heat Transfer over a porous Stretching Sheet with Internal Heat Generation and Stress Work. IJMM, 2 (2007), 267-281
27. Veena,P.H., Abel.S., RajGopal, K and Pravin ,V.K: Heat transfer in a visco-elastic fluid past a stretching sheet with viscous dissipation and internal Heat Generation. ZAMP , 57 (2006) 447-463.
28. Vafai, K and Tien, C.L.: Boundary and Inertia Effects on Flow and Heat Transfer in Porous Media. Int. J. of Heat Mass Transfer, 24(1981); 195-203
29. Sanyal, D.C., and Dasgupta ; S., : steady Flow and Heat Transfer of a conducting fluid by stretching porous wall. Ind.J.Theo. Phys., 51, (2003) 47-58.
30. Elbaashbeshy, E.M.A., and Bazid, M.A.A: Heat transfer over a steady stretching Surface. J. Heat and Mass Transfer, 41. (2004) 1-4.

Failure Rate Analysis of IC Engine Components

Mr. S. Godwin Barnabas¹, N.Sarathkumar², N. Aswinkumar², M. Venkatesh²

¹Assistant professor, Velammal College of Engineering & Technology.

²Student, Velammal College of Engineering & Technology.

ABSTRACT: The main aim of this paper is to analyze the failure of IC engine components. By analyzing the failure rate of the components in IC engines and also to find out failure range for each and every component. For doing, the real time failure data's and their life periods for each components in the IC engines has been analyzed, from these data's the amount of defects in their original production activities and also defects after the design modification work also been concluded. Based on the failure data's the criticality for each component has been ranked out and risk priority number (RPN) and the corresponding transformed scale for each component has been sorted.

Keyword(s): risk priority number, failure range, transformed scale, Design modification.

I. INTRODUCTION

Internal combustion (IC) engine is a complex power generating machines and used widely in automotive industry, which the failure rate is high. Carrying out the IC engine fault diagnostic methods have been studied and still a lasting topic for scientists. Failure rate is the frequency with which a component fails. The failure rate of a system depends on the time, with the rate varying over the life cycle of the system. Failure rate is defined as the total number of failures within an item population divided by the total time expended by that population, during a particular measurement interval under stated conditions. Engine failures result from a complex set of conditions, effects, and situations. To understand why engines fail and remedy to those failures, one must understand how engine components are designed and manufactured, how they function, and how they interact with other engine components. The failure rate is often thought as the probability that occurs in a specified interval before time. Failure is often denoted by the Greek letter λ (lambda) and is important in reliability engineering. In practice, the mean time between failures $1/\lambda$ (MTBF) is often reported instead of the failure rate. If the failure rate is assumed constant, it may be useful. The MTBF is an important system parameter in systems where failure needs to be managed, in particular for safety Systems. The MTBF appears frequently in the engineering design equipments, where the time to recover from failure can be neglected and the failure remains constant with respect to time. It is simply said as failure in the inverse of the MTBF. Failure rates can be expressed using any measure of time but hours is the most common unit in practice.

II. Literature review

Ravindra Prasad et al [1] used a numerical method is presented for calculating the temperature fields in a semi-adiabatic diesel engine piston having a cooling oil canal. The crown face of the piston is coated by a 2 mm thick oxide based ceramic insulating material. The non-ideal

thermal contacts between the piston circumference and cylinder wall are also considered. A detailed analysis has been given for estimating the boundary conditions of the cylinder-piston assembly of an internal combustion engine. The isothermic distribution in the piston body and the heat flow rate through the different cooling media at four different engine loads have been depicted both for the cases with and without insulation coating. The results indicate a reduction (12–30%) in heat loss through the piston by use of an insulation coating at the piston crown face, assuming that both the heat transfer process from and the temperature of the combustion products remain unchanged.

D.J. Pickens [2] in this paper describes the theory and use of a method for estimating the service life of an internal combustion (I.C.) engine based on experimental evidence and the law of adhesive wear. A simple computer program is described, which predicts the overall life of an I.C. engine from its design data and a typical sample of its particular running conditions. The use of the program for an engine generator set operating on biogas at a farm site is given as an example. We are thoroughly implementing the maintenance, inspection, and operation of diesel engines in order to maintain them in optimum working condition. However, despite the remarkable progress in technology, the number of failures in newly built diesel engine has been increasing. Judging from a number of instances, they seem due to design defects, material defects, and manufacturing faults. Once a diesel engine failure occurs, a ship owner not only loses profits, but can also encounter other major problems, such as the loss of life and environmental damage. Over a period of several years (to make clear the actual conditions) we have attempted to gather and accumulate data on failures and on abnormalities in regard to newly built diesel engines from 15 Japanese ship owners/managers. Our investigation shows that most of these failures are attributable to poor engineering design and poor quality control. Because we (ship owners/operators/managers) want to help improve the reliability of these high-powered diesel engines, we are willing to work with engine designers and builders. We will, therefore, based upon our analysis results, make constructive and positive proposals to engine designers and builders to help them eliminate these problems.

V.Macian [3] concluded combustion failure diagnosis techniques for reciprocating internal combustion engines have been developed over the last few years. Nowadays the most usual techniques are based on the crankshaft instantaneous speed or on engine vibrations. These methods, although successfully in use, may be applied only to maintenance tasks or to low and moderate engine speeds. In this paper, a controller for the correction of injection failures is presented. The aim of the algorithm is to ensure that the same quantity of fuel is injected into each one of the cylinders. This governor can be applied to the full operating range of the engine. The injection failure

detection and identification technique is based on the measurement of the turbocharger instantaneous speed and its treatment in the frequency domain. The simulation of the controller shows an effective reduction in the dispersion between cylinders to a level below 2 per cent.

An expert system solves problems using a process that is very similar to the methods used by the human expert. An Expert System is a computer program designed to model the problem solving ability of a human expert (Durkin, 1994) [2]. When compared to a mechanic, an Expert system would present the following advantages: It is always available and anywhere; it is replaceable; it is not perishable; it is consistent in performance and speed; and its cost is affordable. Currently, there are Expert Systems and computerized tools for diagnosing and troubleshooting car faults in which engine faults can also be diagnosed. Some heavy duty vehicles have On Board Diagnostics (OBD). OBD was developed to provide improved, information - rich visibility to complex operation and control mechanisms that many service technicians still treat as black boxes (Barkai, 2001) [3]. When a simple correlation exists between the OBD malfunction data and its root cause, OBD is a useful troubleshooting tool but it provides little assistance in diagnosing more complex situations such as multiple fault codes or inconsistent information (Barkai, 2001) [3].

The mean time between the failures of the Crank case, Connecting rod, Bearing, Cylinder head, Timing gear IC engine components are collected and they are as follows.

S.no	Crank case (Hrs)	Conn-ecting rod (Hrs)	Beari ng (Hrs)	Cylinder head (Hrs)	Timing gear (Hrs)
1)	20.2	20.13	20.1	20.131	21.231
2)	25.31	25.21	25.3	25.243	47.738
3)	30.54	31.33	30.5	30.335	72.363
4)	35.23	36.41	35.7	35.44	101.94
5)	41.62	41.52	40.9	40.556	130.63
6)	46.84	46.63	46.1	45.634	26.342
7)	52.00	51.74	51.3	50.738	51.844
8)	57.12	57.12	56.5	55.846	78.863
9)	62.45	62.93	61.8	60.95	106.13
10)	67.38	68.13	66.9	65.134	135.72
11)	72.14	73.14	72.1	70.24	31.433
12)	78.02	80.25	77.3	75.331	56.936
13)	83.14	86.31	82.5	80.424	84.632
14)	88.33	95.51	87.7	85.533	111.24
15)	93.60	103.7	93.9	90.64	140.86
16)	98.75	110.8	98.1	95.755	36.521
17)	103.9	116.9	104	100.86	62.14
18)	109.1	121.2	110	105.97	90.743
19)	114.5	127.1	116.4	110.18	117.35
20)	125.8	135.4	123.6	115.16	145.9
21)	130.9	140.5	130.7	120.27	40.83
22)	136.0	155.8	139.9	125.38	66.51
23)	143.2	145.6	148.2	130.42	95.83
24)	152.6	150.7	166.3	135.536	123.46
25)	160.7	160.9	180.5	140.74	150.16

The mean times between the failures of the Crank shaft, valve, camshaft, piston, Cam shaft gear of IC engine components are collected and they are as follows.

These tabulation are done for chi square test, this is done for testing the null hypothesis which states that there is no significant difference between the expected and observed result. Test is done for following IC engine components.

- Crankcase
- Connecting rod
- Bearing
- Cylinder head
- Timing gear
- Crankshaft
- Valve
- Camshaft
- Piston
- Camshaft gear
- Piston
- Camshaft gear

s.no	Crank shaft (Hrs)	Valve (Hrs)	Cam shaft (Hrs)	Piston (Hrs)	Cam shaft gear (Hrs)
1)	20.12	20.133	25.345	32.331	20.131
2)	45.653	26.242	52.954	25.248	25.248
3)	70.263	31.336	78.681	32.331	30.361
4)	95.745	36.45	111.484	36.462	35.481
5)	120.463	41.531	142.231	42.634	40.593
6)	25.232	46.542	25.345	53.234	45.684
7)	50.748	52.743	52.954	60.963	50.736
8)	75.381	57.856	78.681	60.963	55.881
9)	100.854	62.931	111.484	65.148	60.994
10)	125.578	66.14	142.231	70.334	65.16
11)	30.345	71.25	30.453	85.774	70.271
12)	55.859	76.364	57.133	92.834	75.384
13)	80.493	82.431	85.734	92.834	80.496
14)	105.948	87.543	117.563	97.965	85.584
15)	130.683	93.634	148.481	103.136	90.676
16)	35.432	98.743	35.564	116.463	95.781
17)	60.963	103.863	62.288	121.574	100.891
18)	85.584	109.943	92.145	121.574	105.941
19)	110.234	115.245	123.641	126.683	110.136
20)	135.791	121.363	155.563	131.743	115.241
21)	40.548	125.474	41.671	141.948	120.374
22)	65.148	132.694	67.361	147.154	125.483
23)	90.631	138.785	98.268	147.154	130.594
24)	115.348	145.836	130.937	153.236	135.684
25)	140.848	150.945	160.648	160.341	140.731

METHODOLOGY CHI SQUARE TEST

From the life time of all the IC engine components shown in the tabulation the chi square test has been conducted to estimate the mean life time of IC engine components. Chi square test is a statistical test commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis. The chi square test is always testing the null hypothesis which states that there is no significant difference between the expected and observed result. Chi square is the sum of the squared

difference between observed (o) and the expected (e) data (or the deviation, d), divided by the expected data in all possible categories. The degrees of freedom are determined by calculating as the number of components. A relative standard is determined as the basis for accepting or rejecting the hypothesis. The relatively standard commonly used is $p > 0.05$ where p is the probability. Chi square should not be calculated if the expected value in any category is less than 5.

Chi square test is given by,
 $[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 Where T= Total time,
 α =confidence level,
 n= number of components,

III. CALCULATION

1. Crankcase:

The confidence level α is taken as 95%. T is the total mean time of IC engine components from the data's collected.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2392.12/\psi^2 54, 0.975; 2*2392.12/\psi^2 54, 0.025]$
 $[4784.24/\psi^2 54, 0.975; 4784.24/\psi^2 54, 0.025]$
 $[4784.24/68.3, 0.0975; 4784.24/73.6, 0.025]$
 $[68.3; 73.6]$
 $[1/68.3; 1/73.6]$
 $[0.0146; 0.0135]$

The failure range of the crankcase is from 0.0135 to 0.0146 months.

2. Connecting rod:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[5639.14/\psi^2 60, 0.975; 5639.14/\psi^2 60, 0.025]$
 $[5639.14/76.2, 0.975; 5639.14/83.3, 0.025]$
 $[76.2; 83.3]$
 $[1/76.2; 1/83.3]$
 $[0.0131; 0.0120]$

The failure range of the connecting rod is from 0.0120 to 0.0131 months.

3. Bearing:

The confidence level is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2624.65/\psi^2 54, 0.975; 2*2624.65/\psi^2 54, 0.025]$
 $[5249.30/\psi^2 54, 0.975; 5249.30/\psi^2 54, 0.025]$
 $[5249.30/73.4, 0.975; 5249.30/81.5, 0.025]$
 $[73.4, 81.5]$
 $[1/73.4; 1/81.5]$
 $[0.0136; 0.0122]$

The failure range of the bearing is from 0.0122 to 0.0136 months

4. Cylinder head:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2012.43/\psi^2 50, 0.975; 2*2012.43/\psi^2 50, 0.025]$
 $[4024.87/\psi^2 50, 0.975; 4024.87/\psi^2 50, 0.025]$
 $[4024.87/63.3, 0.975; 4024.87/71.4, 0.025]$
 $[63.3; 71.4]$
 $[1/63.3; 1/71.4]$

[0.0157; 0.0140]

The failure range of the cylinder is from 0.0140 to 0.0157 months

5. Timing gear:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2309.64/\psi^2 54, 0.975; 2*2309.64/\psi^2 54, 0.025]$
 $[4619.28/\psi^2 54, 0.975; 4619.28/\psi^2 54, 0.025]$
 $[4619.28/68.3, 0.975; 4619.28/75.6, 0.025]$
 $[68.3, 75.6]$
 $[1/68.3, 1/75.6]$
 $[0.0146, 0.0132]$

The failure range of the timing gear is from 0.0132 to 0.0146 months.

6. Crankshaft:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2013.89/\psi^2 50, 0.975; 2*2013.89/\psi^2 50, 0.025]$
 $[4027.78/\psi^2 50, 0.975; 4027.78/\psi^2 50, 0.025]$
 $[4027.78/63.3, 0.975; 4027.78/71.4, 0.025]$
 $[63.3; 71.4]$
 $[1/63.3; 1/71.4]$
 $[0.0157; 0.0140]$

The failure range of the crankshaft is from 0.0140 to 0.015 months.

7. Valve:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2309.83/\psi^2 54, 0.975; 2*2309.83/\psi^2 54, 0.025]$
 $[4619.66/\psi^2 54, 0.975; 4619.66/\psi^2 54, 0.025]$
 $[4619.66/68.3, 0.975; 4619.66/75.6, 0.025]$
 $[68.3; 75.6]$
 $[1/68.3; 1/75.6]$
 $[0.0146; 0.0132]$

The failure range of the valve is from 0.0132 to 0.0146 months.

8. Camshaft:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2624.76/\psi^2 58, 0.975; 2*2624.76/\psi^2 58, 0.025]$
 $[5249.52/\psi^2 58, 0.975; 5249.52/\psi^2 58, 0.025]$
 $[5249.52/73.4, 0.975; 5249.52/81.5, 0.025]$
 $[73.4; 81.5]$
 $[1/73.4; 1/81.5]$
 $[0.0136; 0.0122]$

The failure range of the camshaft is from 0.0122 to 0.0136 months.

9. Piston:

The confidence level α is taken as 95%. T is the total mean time.

$[2T/\psi^2 2n, 1-\alpha/2; 2T/\psi^2 2n, \alpha/2]$
 $[2*2625.30/\psi^2 58, 0.975; 2*2625.30/\psi^2 58, 0.025]$
 $[5250.60/\psi^2 58, 0.975; 5250.60/\psi^2 58, 0.025]$

[5250.60/74.4, 0.975;5250.60/82.6, 0.025]
 [74.4; 82.6]
 [1/74.4;1/82.6]
 [0.0134; 0.0121]

The failure range of the piston is from 0.0121 to 0.0134 months.

10. Camshaft gear:

The confidence level α is taken as 95%. T is the total mean time.

[2T/ $\psi^2 2n$, 1- $\alpha/2$; 2T/ $\psi^2 2n$, $\alpha/2$]
 [2*2013.53/ $\psi^2 50$, 0.975;2*2013.53/ $\psi^2 50$, 0.025]
 [4027.06/ $\psi^2 50$, 0.975;4027.06/ $\psi^2 50$, 0.025]
 [4027.06/73.4, 0.975;4027.06/81.5, 0.025]
 [73.4; 81.5]
 [1/73.4;1/81.5]
 [0.0136; 0.0122]

The failure range of the piston is from 0.0122 to 0.0134 months.

From these entire test conducted the failure rate of the IC engine components are tabulated as follows.

S.no	Component	occurrence	description	Potential failure range	Rank
1	Crankcase	High	Repeated failures	0.0135 to 0.0146	3
2	Connecting rod	Moderate	Occasional failures	0.0120. to 0.0131.	8
3	Bearing	High	Repeated failures	0.0122.to 0.0136	3
4	Cylinder head	High	Repeated failures	0.0140 to 0.0157	1
5	Timing gear	Moderate	Occasional failures	0.0132 to 0.0146.	4
6	Crank shaft	High	Repeated failures	0.0140 to 0.0157	1
7	Valve	High	Repeated failures	0.0132 to 0.0146.	4
8	Cam shaft	Moderate	Occasional failures	0.0122 to 0.0136.	6
9	Piston	High	Repeated failures	0.0121 to 0.0134.	7
10	Camshaft gear	moderate	Occasional failures	0.0122 to 0.0134.	4

IV. FAILURE MODE AND EFFECTS ANALYSIS

FMEA (Failure Modes and Effects Analysis) is used to identify potential failure modes, determine their effects on the operation of the product, and identify actions to mitigate the failures. Design FMEA is methodology for analyzing potential reliability problems early in the design phase where it is possible to take actions to reduce design defects by modification. It is a product design verification activity that can help avoid a large percentage of product design problems before the design is finalized. While anticipating every failure mode is not possible, the development team should formulate a list of potential failure modes as extensively as possible.

A failure mode is the manner by which an equipment or machine failure is observed. It generally describes the way the failure occurs. In FMEA, occurrence is ranked according to the failure probability, which represents the number of failures anticipated during the design life of an item. The range of values and the linguistic terms used to describe the frequency of the failure mode occurrence Failure modes can be observed and represented by occurrence and failure modes can be considered as defects representations of the subsystem (assembly or components). In this paper, we try to find the relationship between occurrence and defects number to estimate the value of k. The aim is to obtain creditable reliability prediction through making good use of design FMEA result, to reduce the time for gathering valid reliability information, and to increase the prediction efficiency.

V. RELIABILITY PREDICTION USING DESIGN SIMILARITY METHOD

New diesel engines are always developed on the basis of existing ones, a great deal of similarities exist between them although there are some variations. Design similarity method utilizes fault rates of existing components to predict fault rates of new products. The failure rate of an existing component can be obtained from sources such as company warranty records, customer maintenance records, component suppliers, or expert elicitation from design or field service engineers. Defects in a component are imperfections that cause inadequacy or failure. The imperfections are always caused in the design and manufacture process. The relationship between failure rate and defect number is expressed as follows:

$$\lambda_o = m * d_o \tag{1}$$

Where λ_o is the failure rate of existing similar components, d_o denotes the total number of known defects, and m is a coefficient.

The failure rate of the new component is calculated as follows:

$$\lambda_n = m * d_n \tag{2}$$

Where λ_n is the failure rate of the new component, d_n is the total

Defects number of the new design:

$$d_o = d_o + d_i - d_e \tag{3}$$

Where d_n is the total number of new defects caused by design modification, d_e is the total number of eliminated defects by design modification.

According to Eq (1), Eq (2) and Eq (3), the failure rate of the new component can be calculated as:

$$\lambda_n = \lambda_o (d_o + d_i - d_e / d_o) \tag{4}$$

The difference between the failure rates of the new and existing products is defined as $\Delta\lambda$, then:

$$\Delta\lambda = \lambda_n - \lambda_o = k\lambda_o \tag{5}$$

Where k represents the coefficient considering the reliability improvement

Because of design modification, then:

$$\lambda_n = \lambda_o - \Delta\lambda = \lambda_o (1 - k) \tag{6}$$

and Eq. (4) can be rewritten as:

$$\lambda_n = \lambda_o (1 - d_e - d_i / d_o) \tag{7}$$

By comparing Eq. (6) and Eq. (7), the relationship between k and defects number is given as follows:

$$k = d_e - d_i / d_o \tag{8}$$

After determining the values of d_o , d_e and d_i the coefficient k can be obtained.

Then the failure rate of the new subsystem/ component can be calculated according to Eq. (7).

After predicting the reliability value of each component, the reliability of the diesel engine system can be estimated on the basis of the reliability block diagram model, which is expressed in Eq. (9):

$$\lambda_s^* = \sum \lambda_i^*$$

where λ_s^* refers to reliability prediction value of the engine system and λ_i^* refers to the reliability value of the its component.

When using design similar method. It is often difficult to obtain defects number exactly in engineering practice. This motivates us to find a relatively feasible method to estimate the defects number.

VI. ESTIMATION k ON THE BASIS OF FMEA:

FMEA (Failure Modes and Effects Analysis) is used to identify potential failure modes, determine their effects on the operation of the product, and identify actions to mitigate the failures. Design FMEA is methodology for analyzing potential reliability problems early in the design phase where it is possible to take actions to reduce design defects by modification. It is a product design verification activity that can help avoid a large percentage of product design problems before the design is finalized. While anticipating every failure mode is not possible, the development team should formulate a list of potential failure modes as extensively as possible. Failure modes can be observed and represented by occurrence, and failure modes can be considered as defects representations of the subsystem (assembly or components). In this work, the relationship between occurrence and defects number to estimate the value of k has been done. The aim is to obtain creditable reliability prediction through making good use of design FMEA result, to reduce the time for gathering valid reliability information, and to increase the prediction efficiency. According to table 1, there exists a nonlinear relationship between failure rate and occurrence rank. It is not possible to produce a linear function of occurrence rank. By multiplying the failure rate by eight, the relationship can be transformed to linear. The transformed scale of failure rate is also shown in table 1. The defects number of existing items is estimated by:

$$d_o = \sum d_j \quad (9)$$

Where d_j is the transformed scale of failure mode occurrence in design FMEA. After design modification, the total number of new defects is given as:

$$d_i = \sum d_t \quad (10)$$

Where d_t is the transformed scale of the i th new failure mode in design FMEA. The eliminated defects number is given as

$$d_e = \sum d_k \quad (11)$$

Where d_k is the transformed scale of k th failure mode in design FMEA. Then the factor k can be calculated.

Case study

A cylinder head gasket is a gasket that sits between the cylinder block and cylinder head in a diesel engine. It is an integral component of the engine and the most critical sealing application in any engine. The cylinder head gasket

must maintain the seal around the combustion chamber at peak operating temperature and pressure. The gasket must seal against air, coolants, combustion and engine oil at their respective peak operating temperature and pressure. The materials used and design employed must be thermally and chemically resistant to the products of combustion and the various chemicals, coolants and oils used in the engine.

In the design process of a new type of diesel engine on the basis of previously used ones, suppose that design modification is made by increasing the flange of cylinder block. The aim is to decrease the occurrence of "Gas leakage" and to reduce the performance degradation probability subsequently. However, the design modification causes a new potential failure mode.

The steps are shown as follows:

(1) Calculate the sum of transformed scales of five failure modes in the previously designed diesel engine:

$$d_o = 0.004 + 0.004 + 0.00005 + 0.00005 + 0.004 = 0.0121$$

(2) Calculate the sum of transformed scales of potential failure modes in the new design:

$$d_i = 0.00005$$

(3) Calculate the sum of transformed scales of eliminated failure modes in the new design:

$$d_e = 0.004$$

Then the factor k can be obtained according to Eq.

$$(8): K = d_e \cdot d_i / d_o = 0.004 \cdot 0.00005 / 0.0121 = 0.3264$$

From the failure range obtained from the chi-square test for each component in the IC engines the transformed scale for each component is listed as follows.

This tabulation is done by considering occurrence in nature.

- Very low
- Low
- Moderate
- High
- Very high

Rank	occurrence	Description	Potential failure rate	Transfor med scale
1	Very low	Failure is unlikely	<1/15xE 5	0.000005
2 3	Low	Relatively few Failures	About 1/1 5xE4 About 1/15xE3	0.00005 0.0005
4 5 6	Moderate	Occasional failures	About 1/2xE3 About 1/4 xE2 About 1/80	0.004 0.02 0.1
7 8	High	Repeated failures	About 1/20 About 1/8	0.4 1.0
9 10	Very High	Failure is almost Inevitable	About 1/3 >1/2	2.7 4.0

Calculation of failure rate of old component and new component:

$$\lambda_0 = m \cdot d_0$$

$$0.1156 = m \cdot 0.0121$$

$$m = 0.1156 / 0.0121$$

$$= 9.553$$

$$d_n = d_o + d_i + d_e$$

$$= 0.0121 + 0.00005 - 0.004$$

$$= 0.00815$$

Where,

d_i =total number of new defects caused by design modification

d_e =total number of eliminated defects by design modification

$$\lambda_n = m \cdot d_n$$

Where,

λ_n = failure rate of new component

d_n = total number of defects in the new design

$$\lambda_n = m \cdot d_n$$

$$= 9.553 \cdot 0.00815$$

$$= 0.0778$$

$$\lambda = \lambda_o(d_o + d_i - d_e / d_o)$$

$$= 0.1156(0.0121 + 0.00005 - 0.004 / 0.0121)$$

$$= 0.0778$$

$$\Delta\lambda = \lambda_o - \lambda_n$$

$$= k\lambda_o$$

$$= 0.3264 \cdot 0.1156$$

$$= 0.0377$$

Where $\Delta\lambda$ = difference between the failure rates of the new and existing products.

MARKOV CHAIN

A markov chain is an order series of states connected by an appropriate transition matrix, a rectangular array in which the elements are transition probabilities which are such that the probability of an event in time period $n+1$ depends only on the state of the system in time period n .

The purpose of using a markov chain is to obtain the failure probabilities for the future.

There is a finite set of states numbered 1, 2... n. The process can be in one, and only one, of these states at a given time are the so-called transition probability P_y , the probability of a transition from state i to state j , is given for every possible combination of i and j , including $i=j$. These transition probabilities are assumed to be stationary (unchanging) over the time period of interest and independent of how state i was reached. Either the initial state in which the process begins is known, or probability distribution of initial states is specified. The transition probabilities P_y can be arranged in the form of what is termed a one-stage stationary transition probability matrix P :

	To			
From	1	2	3 ...	n
1	p_{11}	p_{12}	$p_{13} \dots$	p_{1n}
2	p_{21}	p_{22}	$p_{23} \dots$	p_{2n}
3	p_{31}	p_{32}	$p_{33} \dots$	p_{3n}
n	p_{n1}	p_{n2}	$p_{n3} \dots$	p_{nn}

P is a square matrix with non-negative elements and row elements that sum to unity. Such a matrix is called a stochastic matrix. Any stochastic matrix can serve as a matrix of transition probabilities; together with an initial

probability distribution of states, it completely defines a markov chain.

MARKOV ANALYSIS ALGORITHM

Before we start analyzing a markov process, a problem is presented in which the states of activities are brands of products and transition probabilities represent the likelihood of customers moving from one brand to another. The various steps involved may be summarized as follows:

1. Determine the retention probabilities (groups of customers that do not switch) by dividing the no of failure components retained for the period under review by the total no components of at the beginning of the period.
2. Determine the probabilities associated with the component failures.

(i) Probabilities of component failures can be calculated by dividing the number of components that fail at each period by the number of components manufactured during the period.

(ii) For component failure probabilities, divide the number of has lost by the original number of customers it served.

3. Develop state transition matrix by listing retention probabilities (as calculated in step1) along the main diagonal (upper left to lower right) whereas loss probabilities (calculated in step2) become row values and gain probabilities become column values.

4. Determine the expected future market shares for any period $m-1$ as shown below:

$$[\text{Failure possibilities of period 1}] [\text{State-transition matrix} =$$

$$[\text{Expected component failures in period 2}]$$

$$[\text{Expected component failures in period 2}] [\text{State-transition matrix}] = [\text{Expected component failures in period 3}]$$

$$[\text{Expected component failures in period k-1}] [\text{state transition matrix}]$$

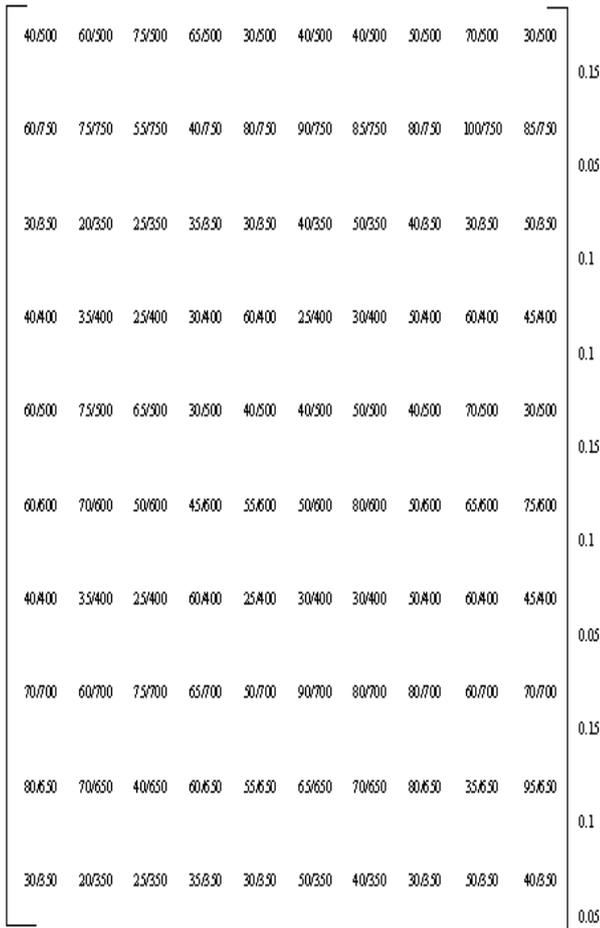
$$= [\text{Expected component failures in period m}]$$

5. Obtain the steady-state or equilibrium conditions for the current problems by the use of matrix algebra and the solution of a set of simultaneous equations obtained above

VII. CALCULATION

$$[\text{Expected component failures in period k-1}] * [\text{state transition matrix}]$$

$$= [\text{Expected component failures in period m}]$$



- X3 no. of failures of bearing (BG)
- X4 no. of failures of Cylinder head (CH)
- X5 no. of failures of timing gear (TG)
- X6 no. of failures of crank shaft (CSH)
- X7 no. of failures of valve (VE)
- X8 no. of failures of camshaft (CMT)
- X9 no. of failures of piston (PN)
- X10 no. of failures of camshaft gear (CG)

$\theta = T/R$
 $\theta = MTBF$
 T = total time
 R = number of failures

by using this relation of all the IC engine components are calculated by the sensitivity analysis conducted on the linear program developed .the sensitivity is conducted by changing the values on the left hand side and also on the right hand side values and also by changing the constraints.

The model linear program is generated from the above relation,

$$\text{Min } x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$$

$$ST$$

$$MTBF1 * x_1 + MTBF2 * x_2 + MTBF3 * x_3 > \sum \text{ of the total life time of the components of CS, CR, BG.}$$

$$MTBF1 * x_1 + MTBF3 * x_3 + MTBF4 * x_4 > \sum \text{ of the total life time of the components of CS, CR, CH.}$$

$$MTBF3 * x_3 + MTBF5 * x_5 + MTBF6 * x_6 > \sum \text{ of the total life}$$

The product of these two matrix provides the upcoming failures of ten components in the IC engines. The following table summarizes the expected failure probabilities for the year 2008 to 2011

VIII. SENSITIVITY ANALYSIS USING LINEAR PROGRAMMING

Sensitivity Analysis for linear Programming model is important, but it is not the only information available. There is a tremendous amount of sensitivity information, or about what happens when data values are changed. We recalled that in order to formulate a problem as a linear program, we had to invoke a certainty Assumption: we had to know what value the data took on, and we made decisions based on that data. Often this assumption is somewhat dubious: the data might be unknown, or guessed. Sensitivity analysis (also called post-optimality analysis) is the study of the behavior of the optimal solution with respect to changes in the input parameters of the original optimization problem. It is often as important solving the original problem itself, partly because in real life applications, the parameters are not always precise and are subject to some source of error. For the LP case, sensitivity analysis based on the optimal basis matrix has been well studied.

Terms used in the sensitivity analysis are as follows:

- X1 no. of failures of crankcase (CS)
- X2 no. of failures of connecting rod (CR)

No.	2008 failure probabilities of 10 IC engine components	2009 failure probabilities of 10 IC engine components	2010 failure probabilities of 10 IC engine components	2011 failure probabilities of 10 IC engine components
1.	0.09905	0.10514	0.0982	0.142
2.	0.10215	0.0984	0.1241	0.0841
3	0.0961	0.1236	0.1091	0.0942
4	0.0915	0.1012	0.0843	0.1041
5	0.085325	0.1082	0.0962	0.0832
6	0.0957	0.1142	0.1241	0.1904
7	0.1213	0.0902	0.1312	0.1014
8	0.10412	0.0854	0.8412	0.0922
9	0.1156	0.1055	0.0804	0.1214
10	0.2594	0.12816	0.0942	0.0734

time of the components of CR,TG,CSH.
 $x_7 > 712$ (total life time of the component of VE)
 $x_8 > 812$ (total life time of the component of CMT)
 $MTBF7 * x_7 + MTBF8 * x_8 + MTBF * x_9 > \sum$ of the total life time of the components of VE, CMT,PN.
 $MTBF6 * x_6 + MTBF7 * x_7 + MTBF8 * x_8 + MTBF9 * x_9 + MTBF10 * x_{10} > \sum$ of the total life time of the components of CSH, VE, CMT, PN, CG.

$$\text{Min } x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$$

$$ST$$

$$40x_1 + 32x_2 + 52x_3 > 2024$$

$$40x_1 + 52x_3 + 20x_4 > 76024$$

$52x_3 + 20x_5 + 23.2x_6 > 1536$
 $x_7 > 512$
 $x_8 > 512$
 $32x_7 + 46x_8 + 56x_9 > 1536$
 $23.2x_6 + 32x_7 + 46x_8 + 56x_9 + 24x_{10} > 2048$ the above framed LP is solved by LINDO and their results are as follows.

OBJECTIVE FUNCTION VALUE

By sensitivity analysis conducted on various IC engine components from X1 to X10 variable cost reductions by comparison is given below. Standard objective functional value is

1) 2486.000

Variable cost	value	Reduced cost
X1	0.000000	0.230769
X2	0.000000	1.000000
X3	1462.000000	0.000000
X4	0.000000	0.615385
X5	0.000000	1.000000
X6	0.000000	1.000000
X7	512.000000	0.000000
X8	512.000000	0.000000
X9	0.000000	1.000000
X10	0.000000	1.000000

Right hand side changes

Min $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$
 ST
 $40x_1 + 32x_2 + 52x_3 > 2124$
 $40x_1 + 52x_3 + 20x_4 > 76324$
 $52x_3 + 20x_5 + 23.2x_6 > 1936$
 $x_7 > 712$
 $x_8 > 812$
 $32x_7 + 46x_8 + 56x_9 > 2036$
 $23.2x_6 + 32x_7 + 46x_8 + 56x_9 + 24x_{10} > 2448$
 END

IX. Results

OBJECTIVE FUNCTION VALUE

1) 1995.467

Left-hand side changes

variable	value	reduced cost
X1	13.62	0.00
X2	0.00	0.72
X3	77.18	0.00
X4	0.00	1.00
X5	44.66	0.00
X6	0.00	0.68
X7	450	0.00

X8	760	0.00
X9	650	0.00
X10	0.00	0.00

Min $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$
 ST
 $45x_1 + 37x_2 + 57x_3 > 2124$
 $35x_1 + 47x_3 + 15x_4 > 76324$
 $55x_3 + 23x_5 + 26.2x_6 > 1936$
 $6x_7 > 712$
 $8x_8 > 812$
 $42x_7 + 56x_8 + 66x_9 > 2036$
 $13.2x_6 + 22x_7 + 36x_8 + 46x_9 + 14x_{10} > 2448$
 END

LP OPTIMUM FOUND AT STEP 0

OBJECTIVE FUNCTION VALUE

1) 1844.082

Changing the constraints

variable	value	reduced cost
X1	0.000000	0.255319
X2	0.000000	1.000000
X3	1623.914917	0.000000
X4	0.000000	0.680851
X5	0.000000	1.000000
X6	0.000000	1.000000
X7	118.666664	0.000000
X8	101.500000	0.000000
X9	0.000000	1.000000
X10	0.000000	1.000000

Max $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$
 ST
 $40x_1 + 32x_2 + 52x_3 < 2024$
 $40x_1 + 52x_3 + 20x_4 < 76024$
 $52x_3 + 20x_5 + 23.2x_6 < 1536$
 $x_7 < 512$
 $x_8 < 512$
 $32x_7 + 46x_8 + 56x_9 < 1536$
 $23.2x_6 + 32x_7 + 46x_8 + 56x_9 + 24x_{10} < 2048$
 END

LP OPTIMUM FOUND AT STEP 4

OBJECTIVE FUNCTION VALUE

1) 4026.583

From the sensitivity analysis conducted on the linear program developed from the data's collected from the IC engine it has been concluded that when the total life time of the components on the right hand side ,MTBF(mean time between the failure) on the left hand side and the inequality constraints are subjected to sensitivity the number of failures becomes minimized by changing the left hand side values compared to changing the values on the values on the right hand side i.e. the total life time of the components .

variable	value	reduced cost
X1	0.000000	2.250000
X2	63.250000	0.000000
X3	0.000000	5.825000
X4	3801.19995	0.000000
X5	76.800003	0.000000
X6	0.000000	1.126667
X7	0.000000	0.333333
X8	0.000000	0.916667
X9	0.000000	1.333333
X10	85.333336	1.000000

X. Conclusion

In this paper from the mean time between the failures of the IC engine components, various failure analyses have been conducted to verify whether the failure rate and failure of the IC engine components are uniform. By the time it is easy to determine the failure range of the IC engine components using chi-square test. In this paper the usage of the markov chain gives the exact failure probabilities of all IC engine components has been determined. The failure mode and effect analysis (FMEA) and cause and effect diagram gives the exact failure reasons, all the design modification problems and finally it prioritizes the IC engines critical components according to their potential failure rate. Finally the sensitivity based optimization is carried out to minimize the total number of failures of the IC engine components.

References

- [1] Ravindra Prasad, N.K. Samria "Investigation of heat transfer in an oil cooled piston with and without ceramic insulation on crown face" International Journal of Mechanical Sciences Volume 31, Issue 10, 1989, Pages 765–777
- [2] D.J. Picken*, H.A. Hassaan* "A method for estimating overhaul life of internal combustion engines including engines operating on biogas and methane" 7 January 1983
- [3] V.Macian "Detection and Correction of injection failures in diesel engines on the basis of turbocharger instantaneous speed frequency analysis" *Universidad Politecnica de valencia, CMT Motores Térmicos Valencia, Spain*.
- [4] Hubert, C. J., Beck, J. W. and Johnson, J. H., "A Model And The Methodology For Determining Wear Particle Generation Rate And Filter Efficiency In A Diesel Engine Using Ferrography", *Wear*, 90 (1983), pp. 335 - 379, (1983).
- [5] . Hargis, S. C., Taylor, H. F. and Gozzo, J. S., "Condition Monitoring Of Marine Diesel Engines Through Ferrographic Oil Analysis", *Wear*, 90 (1983), pp. 225 - 238, (1983).
- [6] Khattab, A. A. and Ali, W. Y., "Development Of Fibrous Oil Filter For Internal Combustion Engines In Desert Environment", *Proceedings of Cairo International Conference On Energy And Environment, Cairo, Egypt, June 3 - 6, (1996)*.
- [7] Balogh, I. and Ali W., "Ferrographic Examination of Solid Particles Contaminating Lubricating Oil", *METALL*, 54, Jahrgang, 4/2000, pp. 129 – 136, (2000).
- [8] Youssef, M., El-Kersh, A. M., Gohar, N. and Ali, W. Y. "Monitoring Automotive Engine Wear By Ferrographic Oil Analysis", *The 5th International Conference of the Egyptian Society of Tribology, Cairo University, EGYPT, 10-12April, (1999)*.
- [9] Khashaba, M. I., Ali, W. Y., and Balogh, I., "Application Of Ferrography In Automotive Engineering", "Lubricant 95", Sopron, Hungary, pp. 103 - 109, (1995).
- [10] Balogh, I. and Ali, W. Y., "Examination of Solid Particles Contaminated in Lubricating Oil by Ferrographic Oil Analysis", *The first International Conference for Advanced Trend in Engineering, Faculty of Engineering, El-Minia University, El-Minia, EGYPT, March 14 - 16, (1999)*

Chemical reaction effect on an unsteady MHD free convection flow past an infinite vertical accelerated plate with constant heat flux, thermal diffusion and diffusion thermo

K. Sudhakar, R. Srinivasa Raju, and M. Rangamma

*Department of Mathematics, University College of Science, Osmania University, Hyderabad, Andhra Pradesh, India.
Department of Mathematics, Padmasri Dr. B. V. Raju Institute of Technology, Narsapur, Medak, Andhra Pradesh, India.*

Abstract: The study of this paper to investigate the effect of chemical reaction on an unsteady magnetohydrodynamic free convection flow of a viscous incompressible fluid past an infinite vertical accelerated plate embedded in porous medium with thermal diffusion, diffusion thermo and constant heat flux in the presence of transverse magnetic field. The governing equations are solved by Galerkin finite element method. The results are obtained for velocity, temperature, concentration, skin friction, Nusselt number and Sherwood number. The effects of different flow parameters on the flow variables are discussed and presented through graphs and tables. And the numerical results for some special cases were compared with Chaudhary et al. [5] and were found to be in good agreement.

Keywords: Thermal diffusion, Diffusion thermo Unsteady, Free convection, MHD, Heat flux, Galerkin finite element method.

Nomenclature:

C	Dimensionless concentration	D	Chemical molecular diffusivity
ε	Porosity of the porous medium	τ_e	Electron collision time in Sec
C'_w	Concentration near the plate	U_o	Reference velocity
C'_∞	Concentration in the fluid far	e	Electron charge, coulombs
θ	Dimensionless Temperature away from the plate	M	Hartmann number
T'	Temperature of the fluid	n_e	Number density of the electron
T'_w	Temperature of the plate	D_m	Mass diffusivity
T'_∞	Temperature of the fluid far away	k_T	Thermal diffusion ratio
	from the plate	c_s	Concentration susceptibility
u'	Velocity component in x' – direction	T_m	Mean fluid temperature
x'	Spatial co – ordinate along the plate	Pr	Prandtl number
ν	Kinematics viscosity, m^2/s plate	P_e	Electron Pressure, N/m^2
y'	Spatial co – ordinate normal to the plate	Sc	Schmidt Number
α	Thermal Diffusivity	g	Acceleration due to Gravity, $9.81 m/s^2$
k_e	Mean absorption coefficient	Gr	Grashof Number
κ	Thermal conductivity, W/mK	k_r	Chemical reaction parameter
σ	Electrical conductivity, mho/m	β	Volumetric co-efficient of thermal Expansion, K^{-1}
μ	Viscosity, Ns/m^2	Gc	Modified Grashof Number
μ_e	Magnetic permeability, Henry/meter	Sr	Soret number
c_p	Specific heat at constant Pressure, J/kg-K	Du	Dufour number
ρ	Density, kg/m^3	β^*	Co – efficient of volume expansion with Species concentration
q'	Radiative heat flux		
ω_e	Electron frequency, radian/sec		

I. Introduction

The phenomenon of hydromagnetic flow with heat and mass transfer in an electrically conducting fluid past a porous plate embedded in a porous medium has attracted the attention of a good number of investigators because of its varied applications in many engineering problems such as MHD generators, plasma studies, nuclear reactors, oil exploration, geothermal energy extractions and in the boundary layer control in the field of aerodynamics. Heat transfer in laminar flow is important in problems dealing with chemical reactions and in dissociating fluids. Combined heat and mass transfer problems with chemical reaction are of importance in many processes and have, therefore, received a considerable amount of attention in recent years. In processes such as drying, evaporation at the surface of a water

body, energy transfer in a wet cooling tower and the flow in a desert cooler, heat and mass transfer occur simultaneously. Possible applications of this type of flow can be found in many industries. For example, in the power industry, among the methods of generating electric power is one in which electrical energy is extracted directly from a moving conducting fluid. Many practical diffusive operations involve the molecular diffusion of a species in the presence of chemical reaction within or at the boundary. There are two types of reactions, homogeneous reaction and heterogeneous reaction. A homogeneous reaction is one that occurs uniformly throughout a given phase. The species generation in a homogeneous reaction is analogous to internal source of heat generation. In contrast, a heterogeneous reaction takes place in a restricted region or within the boundary of a phase. It can therefore be treated as a boundary condition similar to the constant heat flux condition in heat transfer.

The study of heat and mass transfer with chemical reaction is of great practical importance to engineers and scientists because of its almost universal occurrence in many branches of science and engineering. The flow of a fluid past a wedge is of fundamental importance since this type of flow constitutes a general and wide class of flows in which the free stream velocity is proportional to a power of the length coordinate measured from the stagnation point. All industrial chemical processes are designed to transform cheaper raw materials to high value products (usually via chemical reaction). A 'reactor', in which such chemical transformations take place, has to carry out several functions like bringing reactants into intimate contact, providing an appropriate environment (temperature and concentration fields) for adequate time and allowing for removal of products. Fluid dynamics plays a pivotal role in establishing relationship between reactor hardware and reactor performance. For a specific chemistry catalyst, the reactor performance is a complex function of the underlying transport processes. The first step in any reaction engineering analysis is formulating a mathematical framework to describe the rate (and mechanisms) by which one chemical species is converted into another in the absence of any transport limitations (chemical kinetics). Once the intrinsic kinetics is available, the production rate and composition of the products can be related, in principle, to reactor volume, reactor configuration and mode of operation by solving mass, momentum and energy balances over the reactor. This is the central task of a reaction and reactor engineering activity. Analysis of the transport processes and their interaction with chemical reactions can be quite difficult and is intimately connected to the underlying fluid dynamics. Such a combined analysis of chemical and physical processes constitutes the core of chemical reaction engineering. Recent advances in understanding the physics of flows and computational flow modeling (CFM) can make tremendous contributions in chemical engineering.

In view of its wide applications, Acharya *et al.* [1] have reported the problem of heat and mass transfer over an accelerating surface with heat source in presence of suction and blowing. Chamkha and Takhar [3] are used the blotter difference method to study laminar free convection flow of air past a semi infinite vertical plate in the presence of chemical species concentration and thermal radiation effects. Chandran and his associates [4] have discussed the unsteady free convection flow of an electrically conducting fluid with heat flux and accelerated boundary layer motion in presence of a transverse magnetic field. Chaudhary *et al.* [5] studied the effect of free convection effects on magnetohydrodynamic flow past an infinite vertical accelerated plate embedded in porous media with constant heat flux by using Laplace transform technique for finding the analytical solutions. Das and Mitra [6] discussed the unsteady mixed convective MHD flow and mass transfer past an accelerated infinite vertical plate with suction. Recently, Das and his co-workers [7] analyzed the effect of mass transfer on MHD flow and heat transfer past a vertical porous plate through a porous medium under oscillatory suction and heat source. Das *et al.* [8] investigated numerically the unsteady free convective MHD flow past an accelerated vertical plate with suction and heat flux. Das and his associates [9] estimated the mass transfer effects on unsteady flow past an accelerated vertical porous plate with suction employing finite difference analysis.

Gireesh kumar *et al.* [10] investigated effects of chemical reaction and mass transfer on MHD unsteady free convection flow past an infinite vertical plate with constant suction and heat sink. Hasimoto [11] initiated the boundary layer growth on a flat plate with suction or injection. Ibrahim [12] studied the effects of chemical reaction and radiation absorption on transient hydromagnetic natural convection flow with wall transpiration and heat source. Jha [13] analyzed the effect of applied magnetic field on transient free convective flow in a vertical channel. The unsteady free convective MHD flow with heat transfer past a semi-infinite vertical porous moving plate with variable suction has been studied by Kim [14]. Makinde *et al.* [15] discussed the unsteady free convective flow with suction on an accelerating porous plate. Mansutti *et al.* [16] have discussed the steady flow of a non-Newtonian fluid past a porous plate with suction or injection. Sarangi and Jose [18] studied the unsteady free convective MHD flow and mass transfer past a vertical porous plate with variable temperature. Sharma and Pareek [19] explained the behaviour of steady free convective MHD flow past a vertical porous moving surface. Singh and his co-workers [20] have analyzed the effect of heat and mass transfer in MHD flow of a viscous fluid past a vertical plate under oscillatory suction velocity. Singh and Thakur [21] have given an exact solution of a plane unsteady MHD flow of a non-Newtonian fluid. Soundalgekar [22] showed the effect of free convection on steady MHD flow of an electrically conducting fluid past a vertical plate. Yamamoto and Iwamura [23] explained the flow of a viscous fluid with convective acceleration through a porous medium.

Motivate by above reference work, it is proposed here to study the effect of chemical reaction on an unsteady MHD free convection flow past an infinite vertical accelerated plate embedded in porous media with constant heat flux, thermal diffusion and diffusion thermo by Galerkin finite element method which is more economical from computational view point and The results obtained are good agreement with the results of Chaudhary *et al.* [5] in some special cases.

II. Mathematical analysis:

We consider a two – dimensional flow of an incompressible electrically conducting viscous fluid along an infinite non – conducting vertical flat plate through a porous medium. Initially, for time $t' \leq 0$, the plate and the fluid are at some temperature T'_∞ in a stationary condition with the same species concentration C'_∞ at all points. The x' – axis is taken along the plate in the vertically upward direction and the y' – axis is taken normal to the plate. At time $t' > 0$ a magnetic field of uniform strength is applied in the direction of y' – axis and the induced magnetic field is neglected. At time $t' > 0$, the plate starts moving impulsively in its own plane with a velocity U_o with heat supplied to the plate at constant rate. The governing equations of motion and energy under usual Boussinesq’s approximation are given by:

Continuity Equation:

$$\frac{\partial v'}{\partial y'} = 0 \Rightarrow v' = -v'_o \text{ (Constant)} \tag{1}$$

Momentum Equation:

$$\frac{\partial u'}{\partial t'} = \nu \frac{\partial^2 u'}{\partial y'^2} - \frac{\sigma B_o^2 u'}{\rho} + g\beta(T' - T'_\infty) + g\beta^*(C' - C'_\infty) - \frac{\nu u'}{K'} \tag{2}$$

Energy Equation:

$$\frac{\partial T'}{\partial t'} = \frac{\kappa}{\rho c_p} \frac{\partial^2 T'}{\partial y'^2} + \frac{D_m k_T}{c_s c_p} \frac{\partial^2 C'}{\partial y'^2} \tag{3}$$

Diffusion Equation:

$$\frac{\partial C'}{\partial t'} = D \frac{\partial^2 C'}{\partial y'^2} - K'_1 C' + \frac{D_m k_T}{T_m} \frac{\partial^2 T'}{\partial y'^2} \tag{4}$$

With the following initial and boundary conditions:

$$\left. \begin{aligned} t' \leq 0: \{ u' = 0, T' = T'_\infty, C' = C'_\infty \text{ for all } y' \} \\ t' > 0: \left\{ \begin{aligned} u' = U_o, \frac{\partial T'}{\partial y'} = -\frac{q'}{\kappa'}, C' = C'_w \text{ at } y' = 0, \\ u' = 0, T' = T'_\infty, C' = C'_\infty \text{ at } y' \rightarrow \infty \end{aligned} \right\} \end{aligned} \right\} \tag{5}$$

Introducing the following dimensionless quantities:

$$\left. \begin{aligned} t = \frac{t' U_o^2}{\nu}, y = \frac{U_o y'}{\nu}, u = \frac{u'}{U_o}, Pr = \frac{\mu C_p}{\kappa}, Sc = \frac{\nu}{D}, M = \frac{\sigma B_o^2 \nu}{\rho U_o^2}, Gr = \frac{\nu g \beta (T'_w - T'_\infty)}{U_o^3}, \\ Gc = \frac{g \beta^* \nu (C'_w - C'_\infty)}{U_o^3}, K = \frac{U_o^2 K'}{\nu^2}, \theta = \frac{T' - T'_\infty}{T'_w - T'_\infty}, C = \frac{C' - C'_\infty}{C'_w - C'_\infty}, Du = \frac{D_m k_T (C'_w - C'_\infty)}{c_s c_p (T'_w - T'_\infty)}, \\ Sr = \frac{D_m k_T (T'_w - T'_\infty)}{\nu T_m (C'_w - C'_\infty)}, k_r = \frac{K'_1 \nu}{U_o^2} \end{aligned} \right\} \tag{6}$$

Using dimensionless quantities from (6), the equations (2), (3) and (4) reduces to

$$\frac{\partial^2 u}{\partial y^2} + (Gr)\theta + (Gc)C = \frac{\partial u}{\partial t} + (M + \frac{1}{K})u \tag{7}$$

$$\frac{\partial^2 \theta}{\partial y^2} = (Pr) \frac{\partial \theta}{\partial t} - (Pr)(Du) \left(\frac{\partial^2 C}{\partial y^2} \right) \tag{8}$$

$$\frac{\partial^2 C}{\partial y^2} = (Sc) \frac{\partial C}{\partial t} - (Sc)(Sr) \left(\frac{\partial^2 \theta}{\partial y^2} \right) + (Sc)(k_r)C \tag{9}$$

with the following initial and boundary conditions

$$\begin{aligned}
 t \leq 0: & \left\{ u = 0, \theta = 0, C = 0 \text{ for all } y \right. \\
 t > 0: & \left. \begin{cases} u = 1, \frac{d\theta}{dy} = -1, C = 1 \text{ at } y = 0 \\ u = 0, \theta = 0, C = 0 \text{ at } y \rightarrow \infty \end{cases} \right\}
 \end{aligned} \tag{10}$$

All the physical variables are defined in the nomenclature.

III. Method of solution

Applying the Galerkin finite element method for equations (7) – (9) over the element (e) ($y_j \leq y \leq y_k$) yields

$$\int_{y_j}^{y_k} N^{(e)T} \left(\frac{\partial^2 \mathbf{u}^{(e)}}{\partial y^2} - \frac{\partial \mathbf{u}^{(e)}}{\partial t} - B\mathbf{u}^{(e)} + P \right) dy = 0 \tag{11}$$

$$\int_{y_j}^{y_k} N^{(e)T} \left(\frac{\partial^2 \theta^{(e)}}{\partial y^2} - Pr \frac{\partial \theta^{(e)}}{\partial t} + Q \right) dy = 0 \tag{12}$$

$$\int_{y_j}^{y_k} N^{(e)T} \left(\frac{\partial^2 C^{(e)}}{\partial y^2} - Sc \frac{\partial C^{(e)}}{\partial t} - (Sc)(k_r)C^{(e)} + R \right) dy = 0 \tag{13}$$

Where $B = M + \frac{1}{k}$, $P = Gr\theta_i^j + GcC_i^j$, $Q = (Pr)(Du) \left(\frac{\partial^2 C_i^j}{\partial y^2} \right)$, $R = (Sc)(Sr) \left(\frac{\partial^2 \theta_i^j}{\partial y^2} \right)$

Let the linear piecewise approximation solution be

$$\mathbf{u}^{(e)} = N_j(y)u_j(t) + N_k(y)u_k(t) = N_j u_j + N_k u_k,$$

$$\theta^{(e)} = N_j(y)\theta_j(t) + N_k(y)\theta_k(t) = N_j \theta_j + N_k \theta_k,$$

$$C^{(e)} = N_j(y)C_j(t) + N_k(y)C_k(t) = N_j C_j + N_k C_k.$$

Where $N_j = \frac{y_k - y}{y_k - y_j}$, $N_k = \frac{y - y_j}{y_k - y_j}$, $N^{(e)T} = [N_j \quad N_k]^T = \begin{bmatrix} N_j \\ N_k \end{bmatrix}$.

The differential equations (11) – (13) subjected to the boundary conditions (10) are highly non – linear and coupled, and cannot be solved analytically. Therefore, following Bathe [2] and Reddy [17], we use the finite element method to obtain an accurate and efficient solution to the boundary value problem under consideration. The fundamental steps comprising the method are as follows:

Step 1: Discretize the domain into elements:

The whole domain is divided into a finite number of sub – domains. Each sub – domain is termed by a finite element. The collection of the elements is designated to the finite element mesh.

Step 2: Derive the element equations:

The derivation of the finite element equations, i.e., the algebraic equations among the unknown parameters of the finite element approximation, involves constructing the variational formulation of the differential equation, assuming the form of the approximate solution over a typical finite element, and deriving the finite element equations by substituting the approximate solution into the variational formulation.

Step 3: Assemble the element equations:

The algebraic equations obtained are assembled by imposing the inter – element continuity conditions. This yields a large number of algebraic equations, which can constitute the global finite element model governing the whole flow domain.

Step 4: Impose the boundary conditions:

The physical boundary conditions defined in equation (10) are imposed on the assembled equations.

Step 5: Solve the assembled equations:

The final matrix equation can be solved by a direct or indirect (iterative) method. For computational purposes, the coordinate y varies from 0 to $y_{max} = 10$, where y_{max} represents infinity, i.e., external to the momentum, energy, and concentration boundary layers. The whole domain is divided into 100 line elements with the equal width 0.05, and each element has three nodes. Therefore, after assembly of all the element equations, we obtain a matrix of the order 201×201 . This obtained system of equations after assembly of the element equations is non – linear. Therefore, an iterative scheme is used to solve it. The system is linearized by incorporating the known functions. After applying the given boundary conditions only, a system

of 195 equations remains for the solution which has been solved by using the Gauss elimination method. This process is repeated until the desired accuracy of 5×10^{-4} is obtained.

IV. Skin friction and rate of heat and mass transfer

The skin friction, Nusselt number and Sherwood number are important physical parameters for this type of boundary layer flow. The skin friction at the plate, which in the non – dimensional form is given by

$$\tau = \frac{\tau'_w}{\rho U_o \nu} = \left(\frac{\partial u}{\partial y} \right)_{y=0} \quad (14)$$

The rate of heat transfer coefficient, which in the non – dimensional form in terms of the Nusselt number is given by

$$Nu = -x \frac{\left(\frac{\partial T'}{\partial y'} \right)_{y'=0}}{T'_w - T'_\infty} \Rightarrow Nu Re_x^{-1} = - \left(\frac{\partial \theta}{\partial y} \right)_{y=0} \quad (15)$$

The rate of mass transfer coefficient, which in the non – dimensional form in terms of the Sherwood number, is given by

$$Sh = -x \frac{\left(\frac{\partial C'}{\partial y'} \right)_{y'=0}}{C'_w - C'_\infty} \Rightarrow Sh Re_x^{-1} = - \left(\frac{\partial C}{\partial y} \right)_{y=0} \quad (16)$$

Where $Re = \frac{U_o x}{\nu}$ is the local Reynolds number.

V. Results and Discussions

In order to understand the effects of different parameters in the problem, velocity, temperature and concentration profiles, skin friction, Nusselt number and Sherwood number have been discussed by assigning numerical values to various parameters Grashof number (Gr), Modified Grashof number (Gc) Prandtl number (Pr), Schmidt number (Sc), Hartmann number (M), Permeability parameter (K), Soret number (Sr), Dufour number (Du) and Chemical reaction parameter (k_r) separately. We discussed the effects of material parameters on primary velocity profiles from figures (2) to (10), temperature profiles from figures (11) and (12) and concentration profiles from the figures (13) to (15). During the course of numerical calculations of the primary velocity (u), temperature (θ) and concentration (C) the values of the Prandtl number are chosen for Mercury ($Pr = 0.025$), Air at $25^\circ C$ and one atmospheric pressure ($Pr = 0.71$), Water ($Pr = 7.00$) and Water at $4^\circ C$ ($Pr = 11.62$). To focus out attention on numerical values of the results obtained in the study, the values of Sc are chosen for the gases representing diffusing chemical species of most common interest in air namely Hydrogen ($Sc = 0.22$), Helium ($Sc = 0.30$), Water – vapour ($Sc = 0.60$), Oxygen ($Sc = 0.66$) and Ammonia ($Sc = 0.78$). For the physical significance, the numerical discussions in the problem and at $t = 1.0$, stable values for primary velocity, secondary velocity, temperature and concentration fields are obtained. To examine the effect of parameters related to the problem on the velocity field and skin – friction numerical computations are carried out at $Pr = 0.71$. To find out the solution of this problem, we have placed an infinite vertical plate in a finite length in the flow. Hence, we solve the entire problem in a finite boundary. However, in the graphs, the y values vary from 0 to 4, and the velocity, temperature, and concentration tend to zero as y tends to 4. This is true for any value of y . Thus, we have considered finite length.

1.1 Velocity field

The velocity of the flow field is found to change more or less with the variation of the flow of nine parameters. The major factors affecting the velocity of the flow field are Grashof number (Gr), Modified Grashof number (Gc) Prandtl number (Pr), Schmidt number (Sc), Hartmann number (M), Permeability parameter (K), Soret number (Sr), Dufour number (Du) and Chemical reaction parameter (k_r). The effects of these parameters on the velocity field have been analyzed with the help of figures (2) to (10). Figure (2) shows the effect of Grashof number for heat transfer

on velocity. The Grashof number Gr for heat transfer is found to enhance velocity at all points due to the action of free convection current in the flow field. Figure (3) presents the effect of Grashof number for mass transfer on velocity. The figure shows the accelerating effect of the parameter Gc on the velocity of the flow field at all points. In figure (4), we depict the effect of Prandtl number on velocity of the flow field. The presence of heavier Prandtl number in the flow field is found to decelerate velocity at all points. In figure (5) we depict the effect of Schmidt number on velocity of the flow field. The presence of heavier Schmidt number in the flow field is found to decelerate velocity at all points. The effect of Hartmann number M is shown in the figure (6). It is observed that the velocity of the fluid decreases with the increase of Hartmann number values. As expected, the velocity decreases with an increase in the Hartmann number. It is because that the application of transverse magnetic field will result in a resistive type force (Lorentz force) similar to drag force which tends to resist the fluid flow and thus reducing its velocity. Also, the boundary layer thickness decreases with an increase in the Hartmann number. We also see that velocity profiles decrease with the increase of magnetic effect indicating that magnetic field tends to retard the motion of the fluid. Magnetic field may control the flow characteristics. Figure (7) shows the effect of the permeability of the porous medium parameter (K) on the velocity distribution. As shown, the velocity is increasing with the increasing dimensionless porous medium parameter. The effect of the dimensionless porous medium K becomes smaller as K increase. Physically, this result can be achieved when the holes of the porous medium may be neglected. From figures (8) and (9), the effects of Soret and Dufour numbers on the velocity field are shown. We observe that the velocity increases with the increase of both Dufour and Soret number. Figure (10) displays the effect of the chemical reaction parameter (k_r) on the velocity profiles. As expected, the presence of the chemical reaction significantly affects the velocity profiles. It should be mentioned that the studied case is for a destructive chemical reaction (k_r). In fact, as chemical reaction (k_r) increases, the considerable reduction in the velocity profiles is predicted, and the presence of the peak indicates that the maximum value of the velocity occurs in the body of the fluid close to the surface but not at the surface.

5.2 Temperature field

The temperature of the flow field suffers a substantial change with the variation of the flow parameters such as Prandtl number (Pr) and Dufour number (Du). These variations are shown in figures (11) and (12). An increase in Prandtl number decreases the Temperature field (figure (11)). Also, Temperature field falls more rapidly for Water in comparison to Air and the Temperature field curve is exactly linear for Mercury, which is more sensible towards change in Temperature. From this observation it is concluded that Mercury is most effective for maintaining Temperature differences can be used efficiently in the laboratory. Air can replace Mercury, the effectiveness of maintaining the Temperature changes are much less than Mercury. If Temperatures are maintained, Air can be better and cheap replacement for industrial purposes. The Dufour number (Du) does not enter directly into the momentum and mass equations. Thus the effect of Dufour number on velocity and mass profiles is not apparent. Figure (12) shows the variation of temperature profiles for different values of Du . The parameter Du has marked effects on the temperature profiles. It is observed that the temperature profiles increase with the increasing values of Du . It is also observed from this figure that when $Du = 1.0$, that is, when the ratio between temperature and concentration gradient is very small the temperature profile shows its usual trend of gradual decay. As Dufour number Du becomes large the profiles overshoot the uniform temperature close to the boundary.

5.3 Concentration distribution

The concentration of the flow field suffers a substantial change with the variation of the flow parameters such as Schmidt number (Sc), Soret number (Sr) and Chemical reaction parameter (k_r). These variations are shown in figures from (13) to (15). From figure (13), shows that an increase in Schmidt number decreases the concentration field. Also Concentration field falls slowly and steadily for Hydrogen and Helium but falls very rapidly for Oxygen and Ammonia in comparison to Water vapour. Thus Water vapour can be used for maintaining normal Concentration field and Hydrogen can be used for maintaining effective Concentration field. The Soret number (Sr) does not enter directly into the momentum and energy equations. Thus the effect of Soret number on velocity and temperature profiles is not apparent. Figure (14) shows the variation of concentration profiles for different values of Sr . The parameter Sr has marked effects on the concentration profiles. It is observed that the concentration profiles increase with the increasing values of Sr . It is also observed from this figure that when $Sr = 1.0$, that is, when the ratio between concentration and temperature gradient is very small the concentration profile shows its usual trend of gradual decay. As Soret number Sr becomes large the profiles overshoot the uniform concentration close to the boundary. Figure (15) displays the effect of the chemical reaction parameter (k_r) on concentration profiles. As expected, the presence of the chemical reaction significantly affects the concentration profiles. It should be mentioned that the studied case is for a destructive chemical reaction (k_r). In fact, as chemical reaction (k_r) increases, the concentration decreases. It is evident that the increase in the chemical reaction (k_r) significantly alters the concentration boundary layer thickness but does not alter the momentum boundary layers.

5.4 Skin friction and rate of heat and mass transfer

Table – (1) shows the variation of different values $Gr, Gc, Pr, Sc, M, K, Sr, Du,$ and k_r on skin friction (τ). From this table it is concluded that the skin friction (τ) increases as the values of Gr, Gc, K, Sr, Du increase and this behavior is found just reverse with the increase of Pr, Sc, M and k_r .

Table 1. Variation of numerical values of skin friction (τ) for different values of $Gr, Gc, Sc, Pr, M, K, Sr, Du$ and k_r .

Gr	Gc	Pr	Sc	M	K	Sr	Du	k_r	τ
1.0	1.0	0.71	0.22	2.0	1.0	1.0	1.0	1.0	1.5879
2.0	1.0	0.71	0.22	2.0	1.0	1.0	1.0	1.0	1.8742
1.0	2.0	0.71	0.22	2.0	1.0	1.0	1.0	1.0	1.9873
1.0	1.0	7.00	0.22	2.0	1.0	1.0	1.0	1.0	1.2590
1.0	1.0	0.71	0.60	2.0	1.0	1.0	1.0	1.0	1.3586
1.0	1.0	0.71	0.22	4.0	1.0	1.0	1.0	1.0	1.1167
1.0	1.0	0.71	0.22	2.0	2.0	1.0	1.0	1.0	1.6540
1.0	1.0	0.71	0.22	2.0	1.0	2.0	1.0	1.0	1.7412
1.0	1.0	0.71	0.22	2.0	1.0	1.0	2.0	1.0	1.6984
1.0	1.0	0.71	0.22	2.0	1.0	1.0	1.0	2.0	1.3695

Table – (2) shows the variation of Nusselt number (Nu) different values Pr and Du . From this table it is concluded that the Nusselt number (Nu) increases as the value of Du increases and this behavior is found just reverse with the increase of Pr . Table – (3) shows the variation of Sherwood number (Sh) different values Sc, Sr and k_r . From this table it is concluded that Sherwood number (Sh) increase as the value of Sr increase and this behavior is found just reverse with the increase of Sc and k_r .

Table 2. Variation of Nusselt number (Nu) for different values of Pr, Du and λ

Pr	Du	Nu
0.71	1.0	1.2875
7.00	1.0	1.0067
0.71	2.0	1.3481

Table 3. Variation of Sherwood number (Sh) for different values of Sc, Sr, k_r and λ

Sc	Sr	k_r	Sh
0.22	1.0	1.0	1.0598
0.30	1.0	1.0	0.8436
0.22	2.0	1.0	1.2597
0.22	1.0	2.0	0.7694

In order to ascertain the accuracy of the numerical results, the present skin – friction (τ) results are compared with the previous skin – friction (τ^*) results of Chaudhary *et al.* [5] in table – (4). They are found to be in an excellent agreement.

Table 4: Comparison of present Skin – Friction results (τ_1) with the Skin – Friction results (τ_1^*) obtained by Chaudhary *et al.* [5] for different values of Gr, Pr and M

Gr	Pr	M	τ_1	τ_1^*
1.0	0.71	2.0	1.2254	1.2197
2.0	0.71	2.0	1.3592	1.3465
1.0	0.71	2.0	0.9987	0.9954

VI. Conclusions

In this paper, the governing equations for the effect of chemical reaction on an unsteady magnetohydrodynamic free convection flow of a viscous incompressible fluid past an infinite vertical accelerated plate embedded in porous medium with thermal diffusion, diffusion thermo and constant heat flux in the presence of transverse magnetic field has been presented. Employing the highly efficient finite element method, the leading equations are solved numerically. The results illustrate the flow characteristics for the velocity, temperature, concentration, skin friction, Nusselt number and Sherwood number. The conclusions from these results are:

1. It is observed that the velocity (u) of the fluid increases with the increasing of parameters Gr , Gc , K , Sr , Du and decreases with the increasing of parameters Pr , Sc , M and k_r .
2. The fluid temperature increases with the increasing of Du and decreases with the increasing of Pr .
3. The concentration of the fluid increases with the increasing of Sr and decreases with the increasing of Sc and λ .
4. From table (1), it is concluded that the skin friction (τ) increases with the increasing values of Gr , Gc , K , Sr , Du and this behavior is found just reverse with the increasing of Pr , Sc , M and k_r .
5. From table (2), it is concluded that the Nusselt number (Nu) increases with the increasing values of Du and this behavior is found just reverse with the increasing of Pr .
6. From table (3), it is concluded that the Sherwood number (Sh) increases with the increasing values of Sr and this behavior is found just reverse with the increasing of Sc and k_r .
7. On comparing the skin friction (τ) results with the skin friction (τ^*) results of Chaudhary *et al.* [5] it can be seen that they agree very well.

References

- [1] M. Acharya, L. P. Singh and G. C. Dash, Heat and mass transfer over an accelerating surface with heat source in presence of suction and blowing, *Int. J. Engng. Sci.*, 37, 1999, 1 – 18.
- [2] K. J. Bathe, Finite Element Procedures, *Prentice – Hall, New Jersey*, 1996.
- [3] A. J. Chamkha, H. S. Thakhar and V. M. Soundalgekar, Radiation effects on free convection flow past a semi – infinite vertical plate with mass transfer, *Chemical Engineering Journal*, 84, 2001, 335 – 342.
- [4] P. Chandran, N. C. Sacheti and A. K. Singh, Unsteady hydromagnetic free convection flow with heat flux and accelerated boundary motion, *J. Phys. Soc. Japan*, 67, 1998, 124 – 129.
- [5] R. C. Chaudhary, M. C. Goyal, A. Jain, Free convection effects on MHD flow past an infinite vertical accelerated plate embedded in porous media with constant heat flux, *Matematicas Ensenanza Universitaria*, XVII (2), 2009, 73 – 82.
- [6] S. S. Das, and M. Mitra, Unsteady mixed convective MHD flow and mass transfer past an accelerated infinite vertical plate with suction, *Ind. J. Sci. Tech.*, 2 (5), 2009, 18 – 22.
- [7] S. S. Das, A. Satapathy, J. K. Das and J. P. Panda, Mass transfer effects on MHD flow and heat transfer past a vertical porous plate through a porous medium under oscillatory suction and heat source, *Int. J. Heat Mass Transfer*, 52, 2009, 5962 – 5969.
- [8] S. S. Das, A. Satapathy, J. K. Das and S. K. Sahoo, Numerical solution of unsteady free convective MHD flow past an accelerated vertical plate with suction and heat flux, *J. Ultra Sci. Phys. Sci.*, 19(1), 2007, 105 – 112.
- [9] S. S. Das, S. K. Sahoo and G. C. Dash, Numerical solution of mass transfer effects on an unsteady flow past an accelerated vertical porous plate with suction, *Bull. Malays. Math. Sci. Soc.*, 29(1), 2006, 33 – 42.
- [10] J. Gireesh kumar, P. V. Satyanarayana and S. Ramakrishna, Effects of chemical reaction and mass transfer on MHD unsteady free convection flow past an infinite vertical plate with constant suction and heat sink, *J. Ultra Scientist*, 21(3), 2009, 12 – 28 .
- [11] H. Hasimoto, Boundary layer growth on a flat plate with suction or injection, *J. Phys. Soc. Japan*, 12, 1957, 68 – 72.
- [12] F. S. Ibrahim, A. M. Elaiw and A. A. Bakr, Effect of the chemical reaction and radiation absorption on the unsteady MHD free convection flow past a semi infinite vertical permeable moving plate with heat source and suction, *Communications Nonlinear Science Numerical Simulation*, 13, 2008, 1056 – 1066.
- [13] B. K. Jha, Effects of applied magnetic field on transient free convective flow in a vertical channel, *Ind. J. Pure Appl. Math.*, 29(4), 1998, 441 – 445.
- [14] Y. J. Kim, Unsteady MHD convective heat transfer past a semi – infinite vertical porous moving plate with variable suction, *Int. J. Engg. Sci.*, 38, 2008, 833 – 845.
- [15] O. D. Makinde, J. M. Mango and D. M. Theuri, Unsteady free convection flow with suction on an accelerating porous plate, *AMSEJ. Mod. Meas. Cont.*, B 72 (3), 2003, 39 – 46.
- [16] D. Mansutti, G. Pontrelli and K. R. Rajagopal, Steady flows of non-Newtonian fluids past a porous plate with suction or injection, *Int. J. Num. Methods Fluids*, 17, 1993, 927 – 941.
- [17] J. N. Reddy An Introduction to the Finite Element Method, *McGraw – Hill, New York*, 1985.
- [18] K. C. Sarangi and C. B. Jose, Unsteady free convective MHD flow and mass transfer past a vertical porous plate with variable temperature, *Bull. Cal. Math. Soc.*, 97 (2), 2005, 137 – 146.

- [19] P. R. Sharma and D. Pareek, Steady free convection MHD flow past a vertical porous moving surface, *Ind. J. Theo. Phys.*, 50, 2002, 5 – 13.
- [20] A. K. Singh, A. K. Singh and N. P. Singh, Heat and mass transfer in MHD flow of a viscous fluid past a vertical plate under oscillatory suction velocity, *Ind. J. Pure Appl. Math.*, 34(3), 2003, 429 – 442.
- [21] B. Singh and C. Thakur, An exact solution of plane unsteady MHD non – Newtonian fluid flows, *Ind. J. Pure Appl. Math.*, 33(7), 2002, 993 – 1001.
- [22] V. M. Soundalgekar, Free convection effects on steady MHD flow past a vertical porous plate, *J. Fluid Mech.*, 66, 1974, 541 – 551.
- [23] K. Yamamoto and N. Iwamura, Flow with convective acceleration through a porous medium, *Engg. Math.*, 10, 1976, 41 – 54.

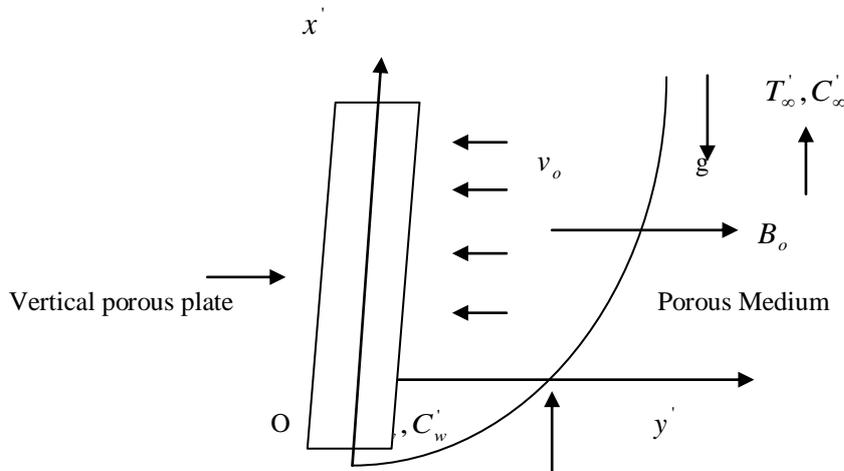


Figure 1. Physical sketch and geometry of the problem

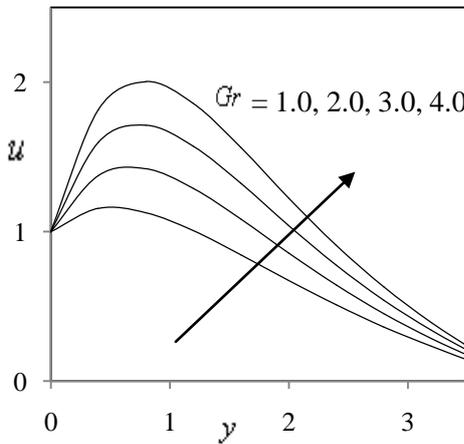


Figure 2. Effect of Grashof number Gr on velocity profiles u

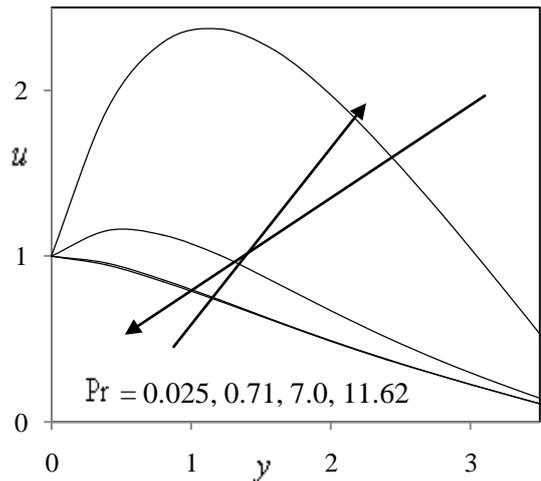


Figure 4. Effect of Prandtl number Pr on velocity profiles u

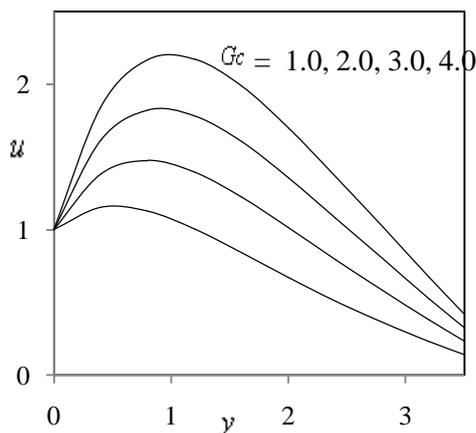


Figure 3. Effect of Modified Grashof number Gc on velocity profiles u

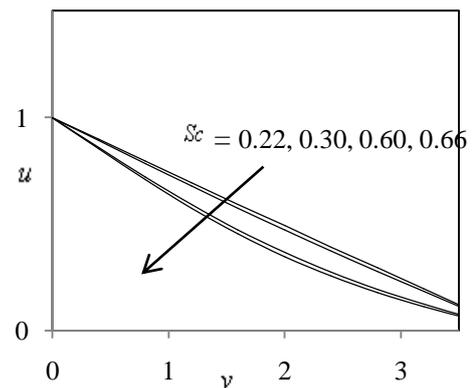


Figure 5. Effect of Schmidt number Sc on velocity profiles u

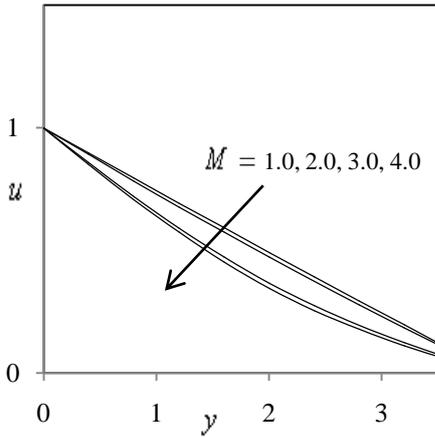


Figure 6. Effect of Hartmann number M on velocity profiles u

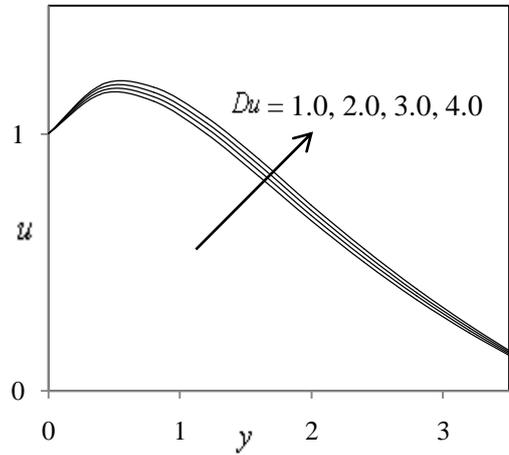


Figure 9. Effect of Dufour number Du on velocity profiles u

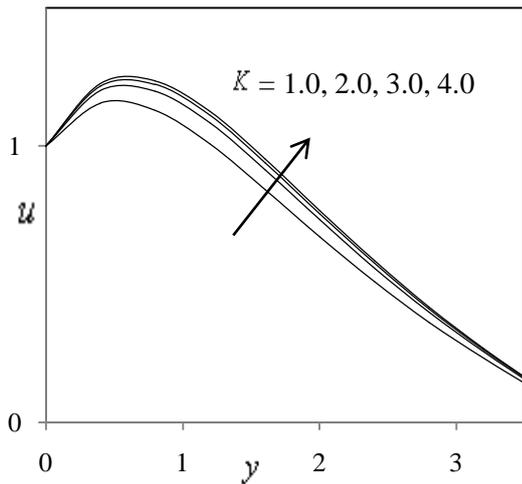


Figure 7. Effect of Permeability parameter K on velocity profiles u

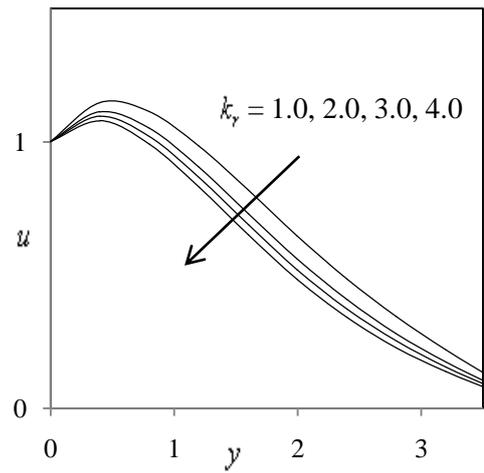


Figure 10. Effect of Chemical reaction parameter k_v on velocity profiles u

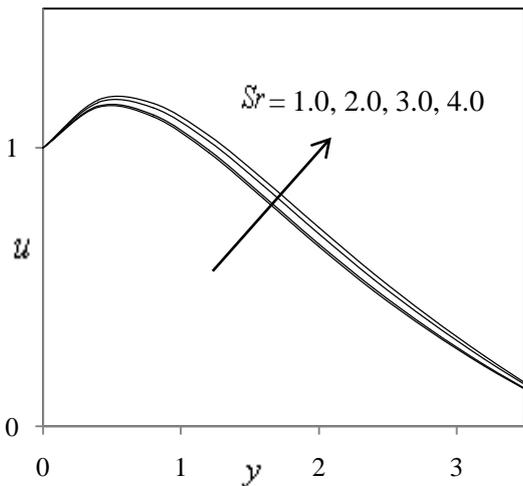


Figure 8. Effect of Soret number Sr on velocity profiles u

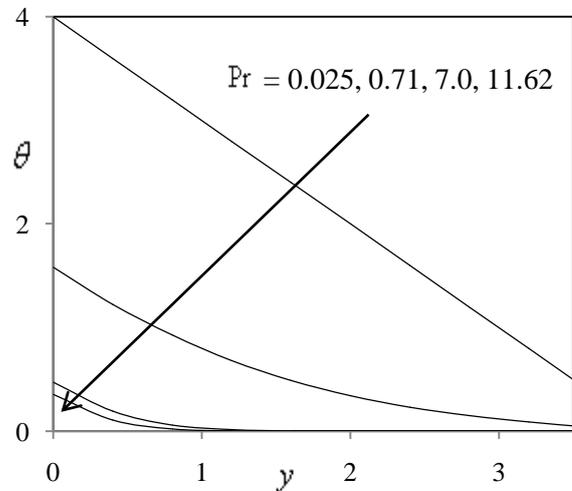


Figure 11. Effect of Prandtl number Pr on temperature profiles θ

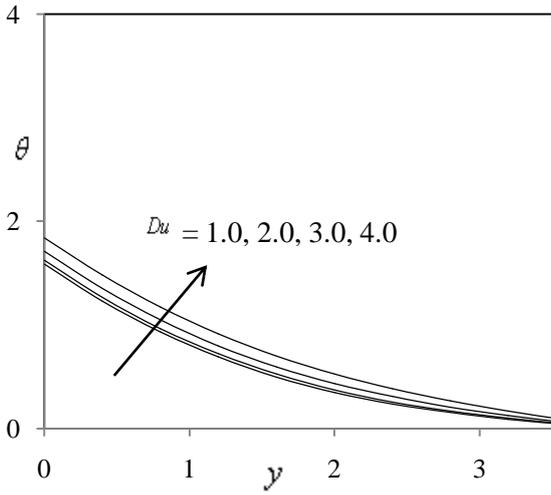


Figure 12. Effect of Dufour number Du on temperature profiles θ

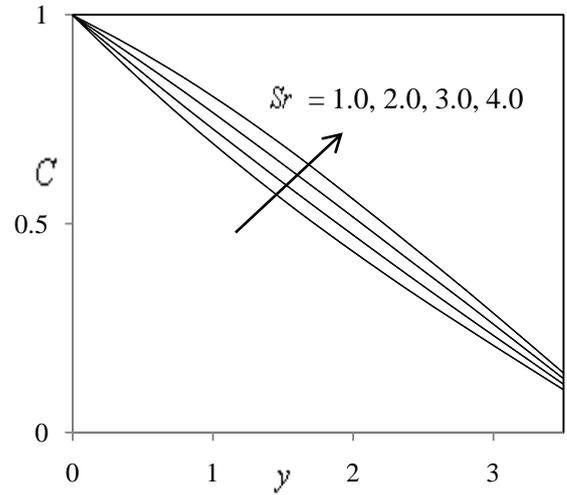


Figure 14. Effect of Soret number Sr on concentration profiles C

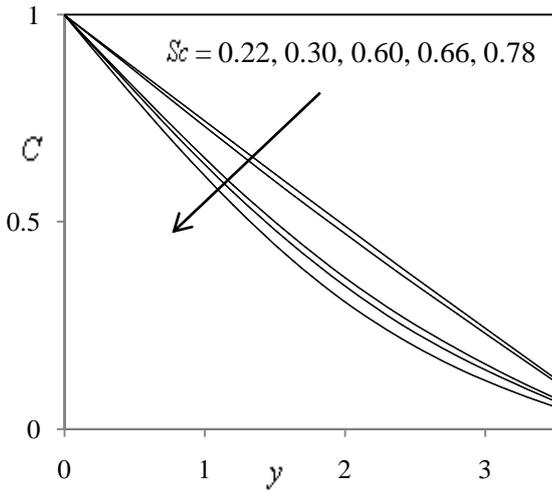


Figure 13. Effect of Schmidt number Sc on concentration profiles C

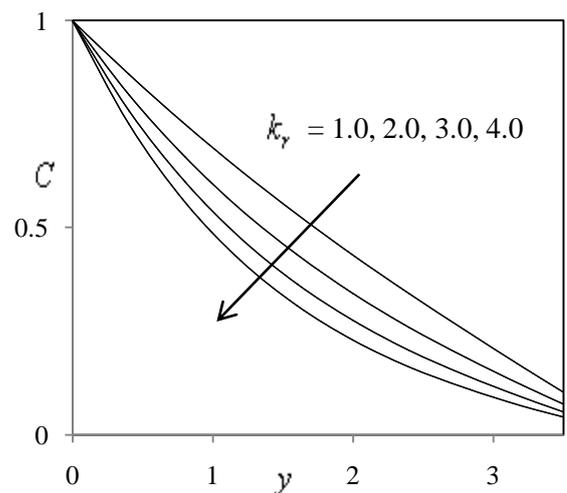


Figure 15. Effect of Chemical reaction parameter k_r on concentration profiles C

A Rational Approach for Fundamental Period of Low and Medium Rise Steel Building Frames

Cinitha. A¹, Umesha. P. K², Nagesh. R. Iyer²

¹ Scientist, CSIR-Structural Engineering research centre, India,

² Chief Scientist, CSIR-Structural Engineering Research Centre, India

³ Director, CSIR-Structural Engineering Research Centre, India

ABSTRACT: The fundamental periods of buildings are essential to calculate the design base shear and lateral forces. Most seismic codes specify empirical formulae to estimate the fundamental vibration period of buildings. These empirical formulae are used for both low- and medium- rise buildings. These formulae depend on the building materials (steel, reinforced concrete etc), building types (frame, shear wall etc) and overall dimensions of the buildings. In this paper, for finding the fundamental natural period of steel moment resisting frames, numerical studies are carried out, and by regression analysis, empirical formulae are derived for low- and medium- rise buildings. The numerical studies carried out includes, the influence of the plan and bay dimensions, normalized stiffness and height of storeys and various cross sections as columns and beams on fundamental period of the moment resisting frames.

Keywords: Dynamic Characteristics, Fundamental Period, Regression Analysis, Normalized stiffness

I. INTRODUCTION

The design of structures to natural hazards such as earthquakes and cyclones demands safety of structures which is governed by the fundamental natural period and the amount of damping in each mode of vibration. Fundamental period of a building and its damping has a remarkable effect on the magnitude of its response. The ability to predict these characteristics, at the design stage, would enable to design safe structures. The fundamental periods of buildings are calculated using the equations specified in building codes for calculating the design base shear and lateral forces. Building codes provide empirical formula that depends on the building material (steel, reinforced concrete etc), building type (frame, shear wall etc) and overall dimensions. Lagomarsino (1993), Tamura (1996), Goel and Chopra (1997) and Tremblay and Rogers (2005) have conducted several field studies on the dynamic characteristics of moment resisting frames. The empirical formula suggested by Goel and Chopra (1997) is adopted in most of the seismic design codes not incorporates, the effect of parameters such as the plan area and symmetry of the building. In this paper, for finding the fundamental natural periods of steel moment resisting frames numerical studies are carried out, and by regression analysis, expressions are derived for low- and medium - rise buildings. The numerical studies carried out includes, the influence of the plan and bay dimensions, normalized stiffness and height of storeys and various cross sections as columns and beams on fundamental period of the moment resisting frames.

II. NUMERICAL MODELING

The numerical studies are carried out with a developed finite element program. The elements used in the program are 3D beam elements with 12 degrees of freedom. The consistent mass matrix is used for analysis. The idealization employs the structure with only one element for each member, which reduces the number of degrees of freedom involved and the computational time. It is assumed that there is no damping in the structure and no time varying forces acts on it. Having defined generalized mass and stiffness matrices, the approximate mode shapes and frequencies of the structure are determined by solving the homogenous equations of the undamped system. The transformation of the mode shapes which result from the solution of the eigenvalue problem in the structural coordinate system to real coordinate system is also accounted. The Cholesky's factorization is used to decompose the stiffness matrix while conducting eigen analysis. Simultaneous iteration method is used to evaluate eigen values and eigen vectors from structural stiffness and mass matrices. The numerical model is validated (Cinitha et al 2008) with experiments conducted on small scale model, Krawinkler et al (1985).

III ESTIMATION OF FUNDAMENTAL PERIODS

The fundamental periods of building frames are highly scattered. The Empirical formulae suggested in building codes are used to calculate the design base shear and lateral forces. Based on their studies Goel and Chopra (1997) suggested an expression as given in equ (1). The fundamental periods found from these expressions are highly conservative. Hence in this study an attempt is made to improve the estimation of periods of steel moment resisting frames without infill. The study is focused on buildings with height varying from 9 to 30m. Studies are carried out with different frame configurations as given in Table.1. For all cases the dimensions of the bay-width and storey-height are assumed to be same. The fundamental natural period is found as per the codal provisions of IS: 1893 (Part 1) 2002 by equ.(2)

$$T = 0.035H^{0.75} \quad (1)$$

$$T_a = 0.085 h^{0.75}, \text{for steel frame building} \quad (2)$$

Where H is the height of the building measured in feet (ft) and h is the height measured in meters.

3.1 NUMERICAL STUDIES AND RESULTS

Numerical studies are carried out to study the effect of the plan and bay dimensions, normalized stiffness and height of storeys of multi bay multi storey frames and various cross section as beams and columns on the fundamental frequency. The ratios of length to width of plan dimensions are taken as plan ratio in the present study.

Fig.1 shows the variation of fundamental frequency with height of the building for various plan dimensions of the building. As reported in empirical formula given in IS:1893-2002, it is found that irrespective of the plan dimensions of the building fundamental frequency decreases with increase in height of the building. The storey stiffness of the building frames are calculated based on the expressions developed by Schultz (1992) and they are normalized according to the maximum structural stiffness observed in each case. Fig.2 shows the fundamental frequency versus normalized stiffness behaviour for buildings with 3x3, 3x4, 3x5 and 3x6 bays. Frames with smaller bays have high fundamental frequency as compared to frames with larger bays. The variation shows a non-linear trend. Fig.3 shows the variation in fundamental frequency with increase in storey height and bay width. A decreasing trend in frequencies is observed with increase in storey height and bay dimensions. As a specific case, frames with different plan ratio are also studied. The fundamental frequency is found for these frames by varying the bay-width from 3m to 6m (3m, 4m, 5m and 6m) and the storey height has been kept constant as 3m. The height of the frame for all cases is assumed as 24m. It is observed that the variation in fundamental frequency is marginal with increase in plan ratio as shown in Fig.4. The buildings with larger bay-width show higher frequency.

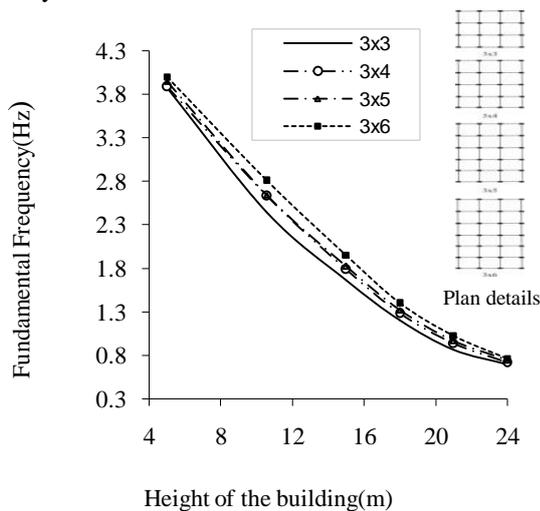


Fig. 1 Variation of fundamental frequency with height of the building

Cases	Parameters varied	
Number of storeys	3, 4, 5, 6 and 7	
Height of Storey (m)	3, 4 and 5	
Spacing of Columns (m)	3, 4, 5, 6 and 9	
Cross Section Details (open sections), (Wide flange sections as per AISC standards)	Beams	Columns
	W14 x 30 , W24 x 62	W16 x 26 , W 18 x 50
Cross Section Details (closed section)	Beams	columns
Case-1	200x200x12	200x200x12
Case-2	300x300x16	300x300x16
Case-3	400x400x20	400x400x20
Case-4	500x500x25	500x500x25

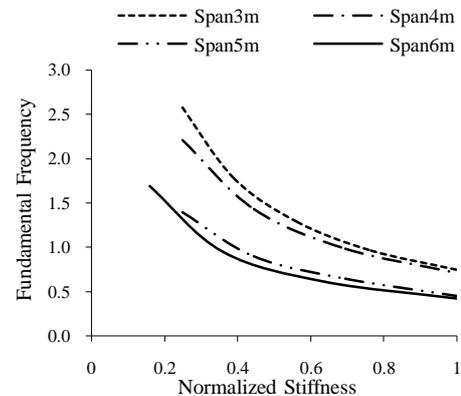


Fig. 2 Variation of fundamental frequency with normalized Stiffness

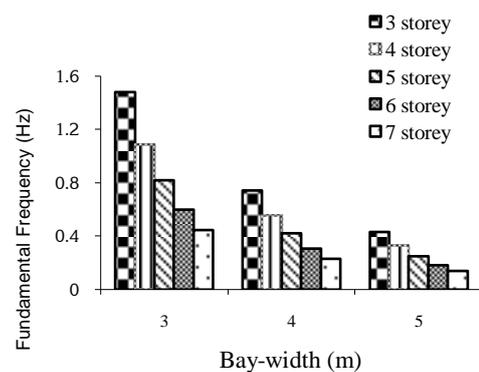


Fig. 3 Fundamental Frequency with Increase in storey height and bay-width

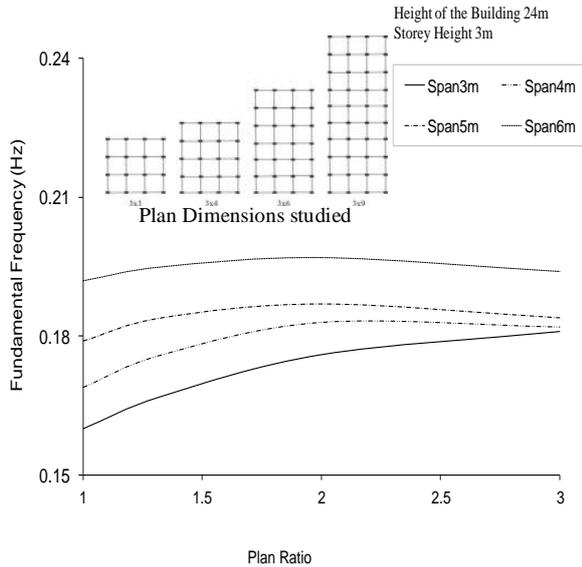


Fig. 4 Fundamental Frequency vs Plan Ratio

III. REGRESSION ANALYSIS

The regular steel moment resisting frames without infill are studied. Based on the height and plan area, parametric studies on 75 regular steel framed structures a regression analysis is carried out for finding their fundamental period. The fundamental periods of the buildings with plan dimensions of 3x3, 3x4, 3x5 and 3x6 (BxD) with heights varied as 3, 4, 5 and 6m, best fit curves are plotted and are shown in Fig.5. In the present study, buildings with a height less than 10m is considered as low-rise buildings and those with 10-30m is considered as medium rise buildings. Among the numerical data, 15 cases with a height of 9m are considered for the prediction of an expression for low rise buildings (0-10m height). The remaining 60 cases are with height of the building as 12, 15, 24 and 30m are considered for the prediction of an expression for medium rise buildings (10-30m, height).

4.1 LOW RISE BUILDINGS

For low rise buildings (i.e., buildings with height less than 10m), from the curve fit the fundamental natural period is derived as a power relation as given in equ.(3) and shown in Fig.6

$$T = 0.056 (BD)^{0.3289} \tag{3}$$

Where, B and D are the length and width of the building plan respectively in 'm'. The present estimation of periods for low-rise buildings is having about 0-13% error as shown in Table.2. Where as the period estimated with the equation suggested in IS: 1893-2002 and Goel and Chopra (1997), the error is found to be 0-77%.

4.2 MEDIUM RISE BUILDINGS

For medium-rise buildings (heights about 10-30m), natural period vs. plan area relationships are derived as power relationships as shown in Fig.7. The variation of the constants with height of the building in power relationships are again fitted to an exponential relationship as shown in the Figs. 8 and 9 respectively. The generalized power relationship in terms of constants C_0 and α is as given in equ.(4). The expressions for C_0 and α value are given

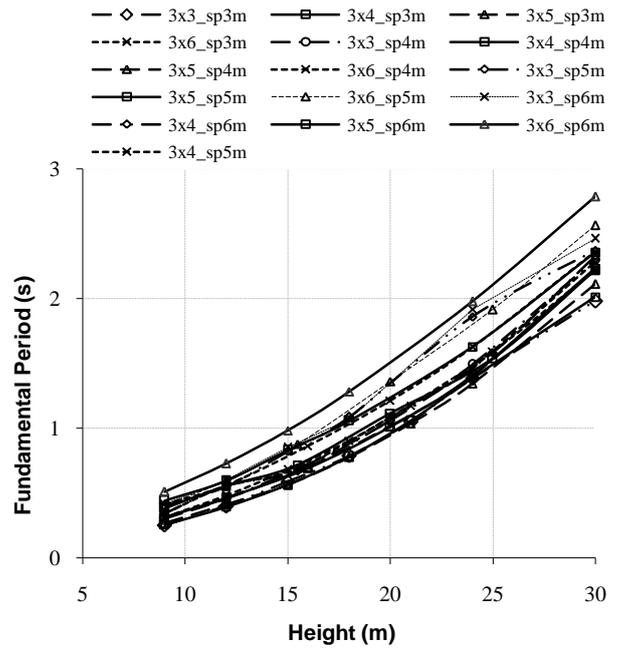


Fig.5 The fundamental Period vs. Height of regular frames studied

in equ. (5) and (6) respectively.

The expression is as follows

$$T = C_0 (BD)^{0.3289\alpha} \tag{4}$$

$$C_0 = 0.0247e^{0.1305 H} \tag{5}$$

$$\alpha = 0.4773e^{-0.0441 H} \tag{6}$$

Where, H is height, Length (B) and width (D) of the building respectively in 'm'. The present estimation of periods for medium – rise buildings is having 0-23% error as shown in Table 3. Where as the period estimated with the equation suggested in IS: 1893: 2002 and Goel and Chopra (1997), the error is found to be about 0-61% as shown in Table 4.

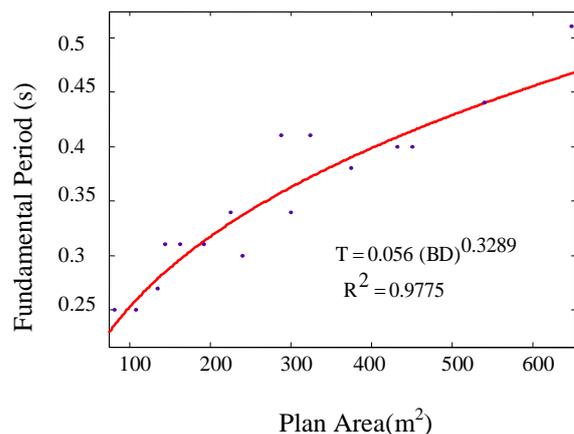


Fig.6 Fundamental Period vs. Plan Area for Low-Rise Buildings

Table 2. Time period (s) for Low- Rise buildings < 10m height, Equ.(3)

Numerical study	Present study (equ.3)	%error	IS:1893 (equ.2)	% error
0.25	0.24	-5.63	0.44	76.67
0.25	0.26	3.74	0.44	76.67
0.27	0.28	3.37	0.44	63.58
0.31	0.30	-4.41	0.44	42.48
0.31	0.29	-8.04	0.44	42.48
0.31	0.31	1.09	0.44	42.48
0.30	0.34	12.41	0.44	47.22
0.41	0.36	-12.67	0.44	7.73
0.34	0.33	-2.90	0.44	29.90
0.34	0.36	6.74	0.44	29.90
0.38	0.39	2.78	0.44	16.23
0.40	0.41	3.67	0.44	10.42
0.41	0.37	-9.22	0.44	7.73
0.40	0.41	2.29	0.44	10.42
0.44	0.44	0.07	0.44	0.38
0.51	0.47	-8.33	0.44	-13.40

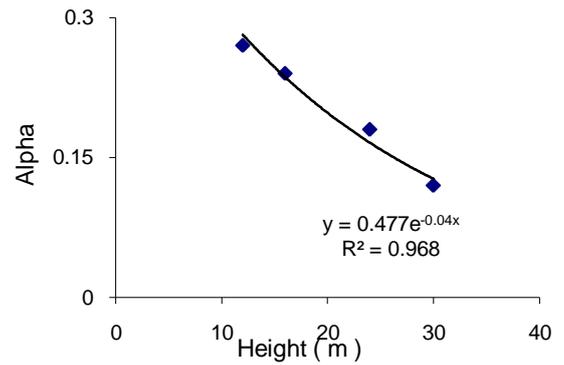


Fig. 8 Alpha vs Height

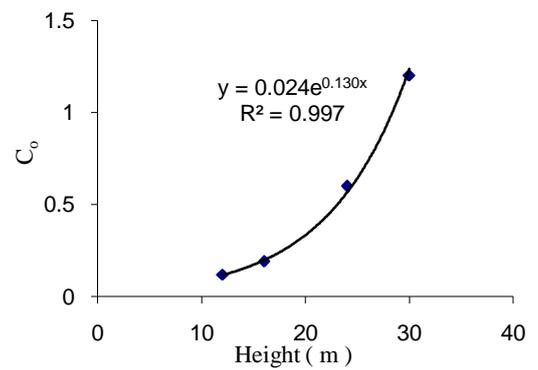


Fig.9 C_o vs. Height

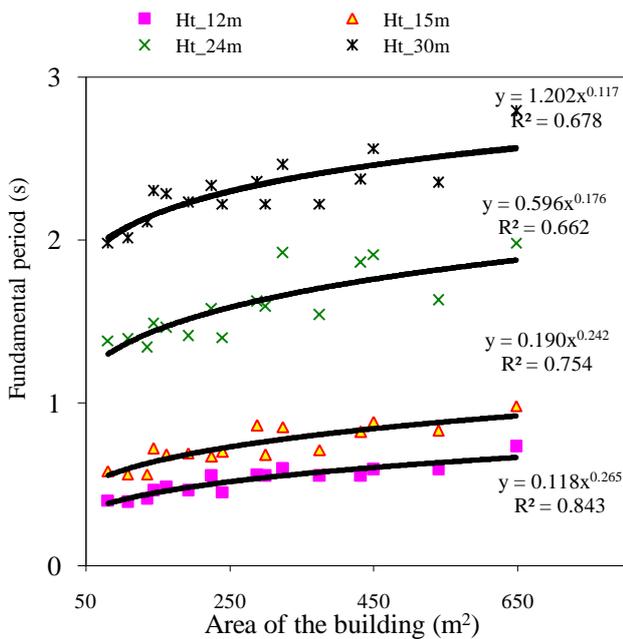


Fig.7 Fundamental Period vs. Plan Area for Medium-Rise Buildings

Area/Height (m)	Numerical Study				Present Study				% error			
	12	15	24	30	12	15	24	30	12	15	24	30
81	0.400	0.581	1.381	1.98	0.41	0.52	1.17	2.17	1.705	10.98	15.06	9.37
108	0.390	0.561	1.392	2.01	0.44	0.55	1.23	2.25	13.1	1.03	11.55	11.75
135	0.410	0.561	1.342	2.11	0.47	0.59	1.28	2.31	14.55	-4.57	4.80	9.52
162	0.480	0.681	1.462	2.28	0.49	0.61	1.31	2.37	2.991	9.93	9.94	3.73
144	0.460	0.721	1.492	2.30	0.48	0.59	1.29	2.33	3.969	17.37	13.46	1.30
192	0.460	0.691	1.412	2.23	0.52	0.64	1.35	2.42	12.73	7.44	4.09	8.37
240	0.450	0.701	1.402	2.22	0.55	0.67	1.40	2.49	22.69	3.61	-0.23	11.99
288	0.560	0.861	1.622	2.36	0.58	0.71	1.45	2.54	3.78	17.94	10.72	7.82
225	0.550	0.671	1.582	2.33	0.54	0.66	1.39	2.47	-1.42	0.88	12.13	5.83
300	0.550	0.681	1.592	2.22	0.59	0.71	1.46	2.56	6.886	-4.83	8.42	15.21
375	0.550	0.711	1.542	2.22	0.63	0.75	1.51	2.63	13.81	-6.08	1.89	18.53
450	0.590	0.881	1.912	2.56	0.66	0.79	1.56	2.69	11.67	10.49	18.47	5.20
324	0.600	0.851	1.922	2.46	0.6	0.73	1.47	2.58	0.122	14.53	23.19	5.00
432	0.550	0.821	1.862	2.37	0.65	0.78	1.55	2.68	18.43	4.90	16.84	13.04
540	0.590	0.831	1.632	2.35	0.69	0.82	1.60	2.76	17.55	0.73	1.54	17.28
648	0.730	0.981	1.982	2.79	0.73	0.86	1.65	2.82	-0	12.07	16.46	1.10

Table 4. Time period in 's' for Medium- Rise buildings with IS:1893 Equ.(2)

Area/ Height (m)	Numerical Study				IS:1893as given in equ2				% error			
	12	15	24	30	12	15	24	30	12	15	24	30
81	0.400	0.581	1.381	1.98	0.55	0.65	0.92	1.09	37.01	11.70	33.21	44.97
108	0.390	0.561	1.392	2.01	0.55	0.65	0.92	1.09	40.52	15.69	33.69	45.79
135	0.410	0.561	1.342	2.11	0.55	0.65	0.92	1.09	33.67	15.69	31.22	48.36
162	0.480	0.681	1.462	2.28	0.55	0.65	0.92	1.09	14.17	-4.73	36.87	52.21
144	0.460	0.721	1.492	2.30	0.55	0.65	0.92	1.09	19.14	-10.0	38.14	52.63
192	0.460	0.691	1.412	2.23	0.55	0.65	0.92	1.09	19.14	-6.11	34.63	51.14
240	0.450	0.701	1.402	2.22	0.55	0.65	0.92	1.09	21.78	-7.45	34.17	50.92
288	0.560	0.861	1.622	2.36	0.55	0.65	0.92	1.09	-2.14	-24.7	43.11	53.83
225	0.550	0.671	1.582	2.33	0.55	0.65	0.92	1.09	0.00	-3.30	41.67	53.24
300	0.550	0.681	1.592	2.22	0.55	0.65	0.92	1.09	0.00	-4.73	42.03	50.92
375	0.550	0.711	1.542	2.22	0.55	0.65	0.92	1.09	0.00	-8.75	40.15	50.92
450	0.590	0.881	1.912	2.56	0.55	0.65	0.92	1.09	-7.11	-26.4	51.74	57.44
324	0.600	0.851	1.922	2.46	0.55	0.65	0.92	1.09	-8.66	-23.8	52.00	55.71
432	0.550	0.821	1.862	2.37	0.55	0.65	0.92	1.09	0.00	-21.0	50.45	54.03
540	0.590	0.831	1.632	2.35	0.55	0.65	0.92	1.09	-7.11	-21.9	43.46	53.63
648	0.730	0.981	1.982	2.79	0.55	0.65	0.92	1.09	-24.9	-33.9	53.45	60.95

Fig 10 and Fig.11 show the fundamental period versus height of the building with closed and open sections respectively, for various plan dimensions of the building. Irrespective of the plan dimensions of the building and irrespective of cross sections fundamental period increases with increase in height of the building. This increase is steep for lighter sections and shallow for heavier sections. It was observed that buildings with heavy sections have

lower fundamental period as compared to buildings with lighter sections. A similar behaviour was observed with frames with open sections.

Figs 12 (a) to (d) show the comparison of fundamental period for buildings with different plan dimensions for various cross sections of beams and columns. For buildings with lower height there is no significant change in fundamental period irrespective of the cross section of beams and column, whereas for buildings with height more than 20m, fundamental period decreases for buildings rectangular in plan. This decrease in

fundamental period is not very significant for buildings with heavier cross section.

IV. RESULTS AND DISCUSSIONS

Based on the numerical studies an expression for finding the fundamental natural periods of low- and medium-rise building is derived. The expression suggested for low-rise buildings are found to be closer to numerical studies than the equation suggested in IS: 1893-2002. The fundamental period predicted with the suggested expression is found to be 0-13% error for low rise building frames compared to IS: 1893-2002 provisions, where the error is 0-77%. But for medium-rise buildings the error is 23% and this can be further improved with analysis of large sampling data.

In general considerable amount of uncertainties exist in the estimation of frequency of the structures by different methods. This depends on the two parameters mass and stiffness, whereas the mass can be predicted fairly accurately. The stiffness in real structure becomes unpredictable because of modeling accuracy of the boundary conditions, materials used in construction and its characterizations etc. Hence, detailed numerical studies were conducted to study the effect of height of the storeys, normalized stiffness of the building and plan and bay dimensions and various cross sections as columns and beams on the fundamental period of moment resisting steel space frames. It is found that the fundamental natural frequency decreases with increase in height of the building irrespective of plan dimensions of the building. The fundamental natural frequencies of frames are decreasing with increase in normalized stiffness and bay-width. It is also found that the variation in fundamental natural frequency is not much significant with increase in plan ratio, but the buildings with larger bay-width have higher frequency. Irrespective of the plan dimensions of the building and irrespective of cross sections fundamental period increases with increase in height of the building. This increase is steep for lighter sections and shallow for heavier sections. It was observed that buildings with heavy sections have lower fundamental period as

compared to buildings with lighter sections. A similar behaviour was observed with frames with open sections.

V. CONCLUSIONS

In this paper based on numerical studies separate expressions are suggested for fundamental natural period of low- and medium- rise buildings. It is found that the fundamental natural frequency decreases with increase in height and normalized stiffness of the building irrespective of plan dimensions of the building. But with increase in plan area the fundamental frequency of the buildings is found to be increasing. It was also observed that buildings with heavy sections as beams and columns have lower fundamental period as compared to buildings with lighter sections.

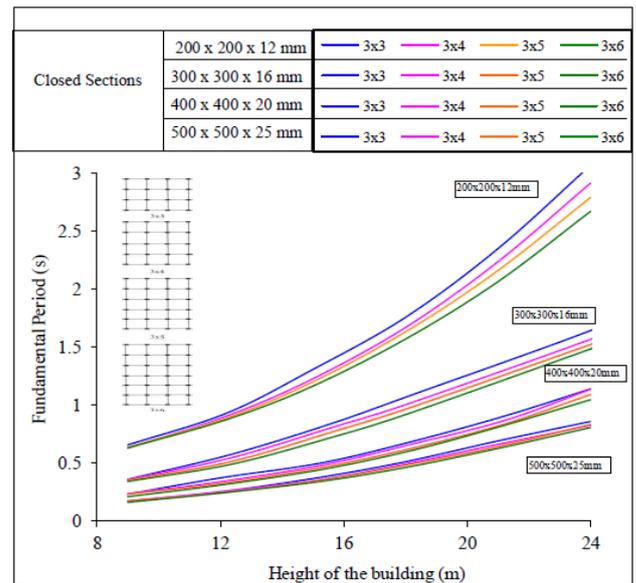


Fig.10 Fundamental period versus Height of the building with closed sections

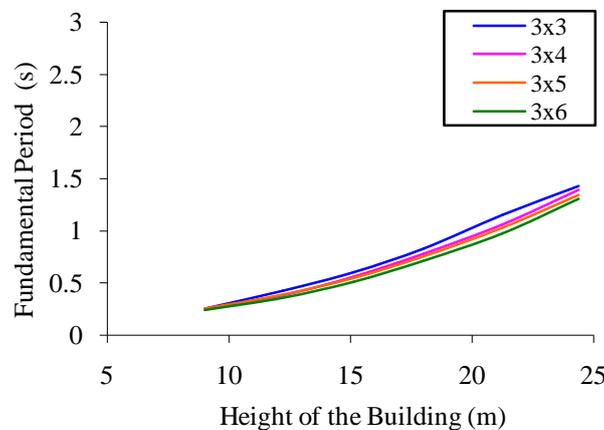


Fig.11 Fundamental period versus Height of the building with Open sections

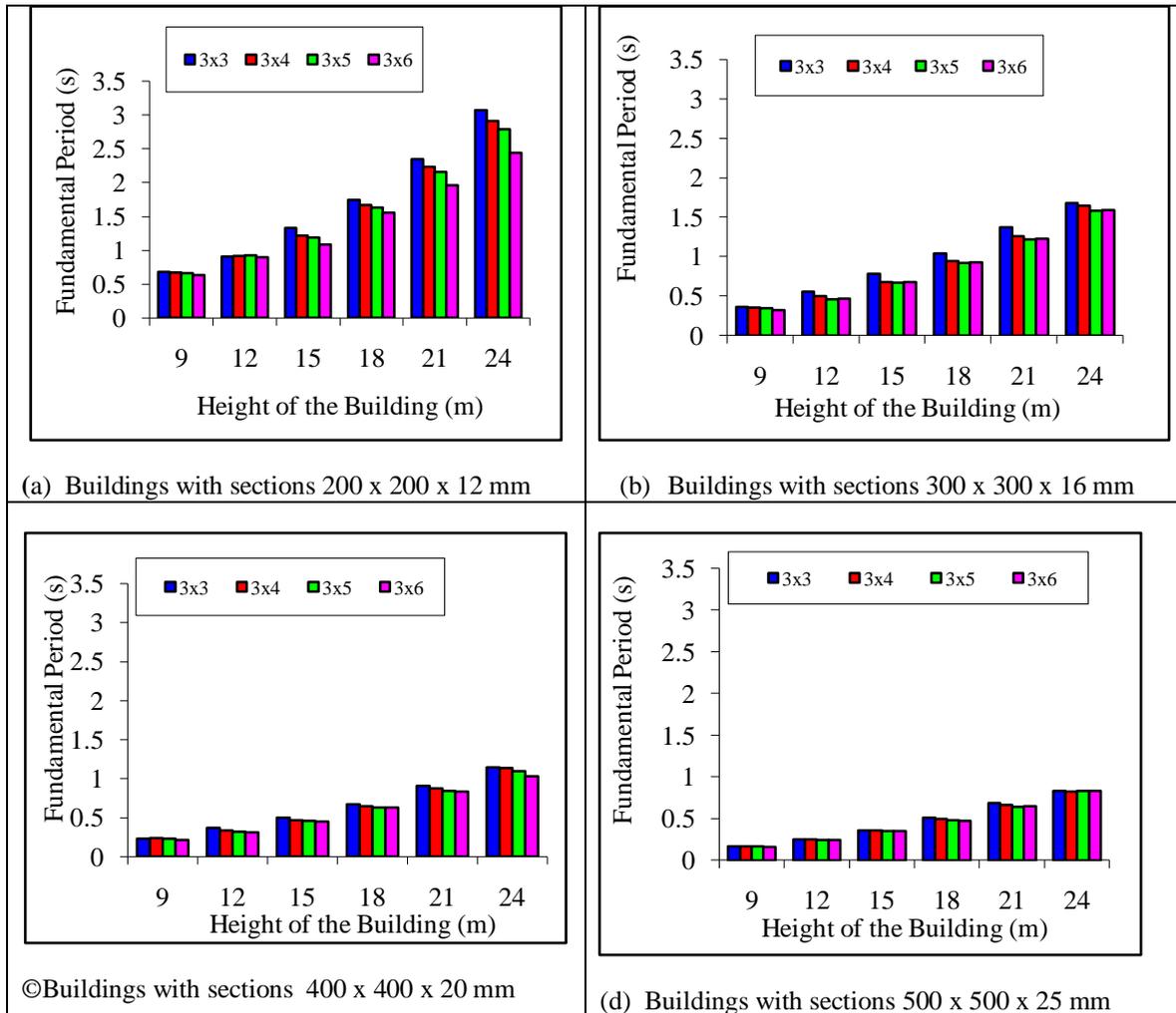


Fig. 12 Comparison of fundamental period for different plan dimensions of the buildings

ACKNOWLEDGEMENT

The authors are grateful to the Director, CSIR-Structural Engineering Research Centre for his encouragement, guidance and suggestions during this research work. The technical discussions with Dr.K.Balaji Rao, Chief Scientist, CSIR-SERC is profoundly thanked. This paper is being published with the permission of the Director, CSIR-Structural Engineering Research Centre.

REFERENCES

- [1] IS 1893 Part I (2002) Criteria for earthquake resistant design of structures
- [2] Cinitha.A., Samuel Knight,G.M., and Ramamurthi.V, Evaluation of Fundamental Frequency of Steel Building Frames-An Experimental Modal Analysis, Journal of the Institution of Engineers,Vol.88, 2008,18-22.
- [3] Goel, K. R.and Chopra, K.A., Period formulas for moment-resisting frame Buildings, J. of Struct. Eng., ASCE, Vol.123, 1997,1454-1461.
- [4] Krawinkler,H. and Benjamin.J. Wallace., Small-scale model experimentation on steel assemblies, Report No.75, 1985, The John A. Blume Earthquake Engineering Centre, Department of Civil Engineering, Stanford University, Stanford.
- [5] Lagomarsino, S. , Forecast models for damping and vibration periods of buildings, J. of Wind Eng. and Ind Aerodyn., Vol. 48, 1993,221-239.
- [6] Schultz A.E., Approximating Lateral Stiffness of Storeys in Elastic Frames, J. of Struct. Eng., ASCE, Vol. 118(1),1992, 243-263.
- [7] Tamura, Y., Sukanuma, S., “Evaluation of amplitude-dependent damping and natural frequency of buildings during strong winds.”, J. of Wind Eng. and Ind. Aerodyn., Vol. 59, 1996, 115-130.
- [8] Tremblay, R. and Rogers, C.A. , Impact of capacity design provisions and period limitations on the seismic design of low-rise steel buildings, Intl. J.of Steel Struct., Vol. 5, 2005 , 1-22.

The Essential Identified Consumer Requirements Derived through Descriptive Analysis On the Information Security in a Smart Grid System

Amy Poh Ai Ling¹, Chen Yan Yu¹, Sugihara Kokichi¹, Mukaidono Masao²

¹Department of Mathematical Modeling Analysis, and Simulation, Graduate School of Advanced Mathematical Sciences, Meiji University, Kanagawa, Japan.

²Computer Science Department, School of Science and Technology, Meiji University, Kanagawa, Japan.

Abstract: The purpose of this paper was to measure the essentiality of the identified consumer requirements on the information security in a smart grid system, focuses mainly on helping government and utilities providers, when enhancing their system processes, to develop and prioritize the necessary requirements for securing their information in a smart grid. This paper examined on the voice of customer perceived by the respondents towards smart grid. The major finding of this paper exhibits the agreement from the respondent on the essentiality of the identified sixteen consumer requirements and the correlation between each requirement. The result of the ranking was a mixture of from five categories of philosophy, human behavior, rule based social system, strategic system and hardware, and hence each category could be concluded to have a balance value of essentiality. Chi-Square test proved that all variables are significant. Individual value plot versus country shows that the rapidly growing economies in different countries have different smart grid infrastructure needs. Pearson correlation test showed that security in wireless media has very low correlation with others variable as it is still in its development process. Lexicographical order test showed that integrity, quality assurance and high bandwidth of communications channels has high correlation with each other. Correlation relationship ring, cluster and tree concluded that the items in the hardware category are highly correlated to each other. The success of this effort appeared to hinge on utility companies championing information system security initiatives and propagating an awareness of the importance of information security among consumers at all levels of the community.

Keywords: smart grid, information security, consumer requirement, descriptive analysis, Pearson correlation, correlation relationship ring.

I. INTRODUCTION

One reason for putting a magnifier over the security term in an information pool of the smart grid system was that cyber technology depends much on its control management to create certain assurance on its protected information stored. Another reason, perhaps more important than the first, was that the security allows the consumer to gain trust towards the service they utilized. For example, many people surfing the net to find out what information will be uploaded to the grid once they become a smart grid user. Others may concern only the benefit a smart grid system brings. Those wanting to go deeper may look at the entire system, its contribution towards environment, cost saving feasibility, and so on. Smart grid has been recognizes as a growing potential and sustainable solution for energy issue. A system will grow strong when it continuously gained its consumer trust and support, thus consumer is one of the key element. In order to realize the smart grid, data were gathered from large numbers of intelligent sensors and processors installed on the power lines and equipment on the distribution grid. This data collected will be transferred to central information processing systems, which both present the information to operators and use the information to send back control settings. While information and communication seems to be much more important compare to decades ago, the system has to be incredibly strong to protect itself against hackers' attacks. The objective of this study was to measure the essentiality of information security requirement based on the consumer perceived significance and the correlation between each requirement.

II. Literature Review

2.1 Technological transfer phenomenon – Grid technology

Modeling lifestyle effects on energy demand, the increase of societal energy consumption was influenced by three main items: technical efficiency, lifestyles and socio-cultural factors. Technological transfer phenomenon was often seen as a crucial part that contributes to the solutions of environmental highlights. However, when technological change was seen from the perspective of everyday life, this image becomes more complex [1, 2]. Grid technology provides the chance for a simple and transparent access to different information sources. This idea was proven when the data grid can be interpreted as the consolidation of different data managing systems furnishing the user with data, information and knowledge [3]. With the changes of smart grid and other in the electricity utility industries, new demands on the telecom networks were generated [4].

2.2 Consumer's impact towards the information security system

When the information was shared real-time between power generators, distributed resources, service provider, control center, substation, and even to end-users, any changes exposed to a hacker's attack would bring the whole system down into a mess. This will dangerously create consumer distrust and dissatisfaction that may lead to other more destructive phenomenon.

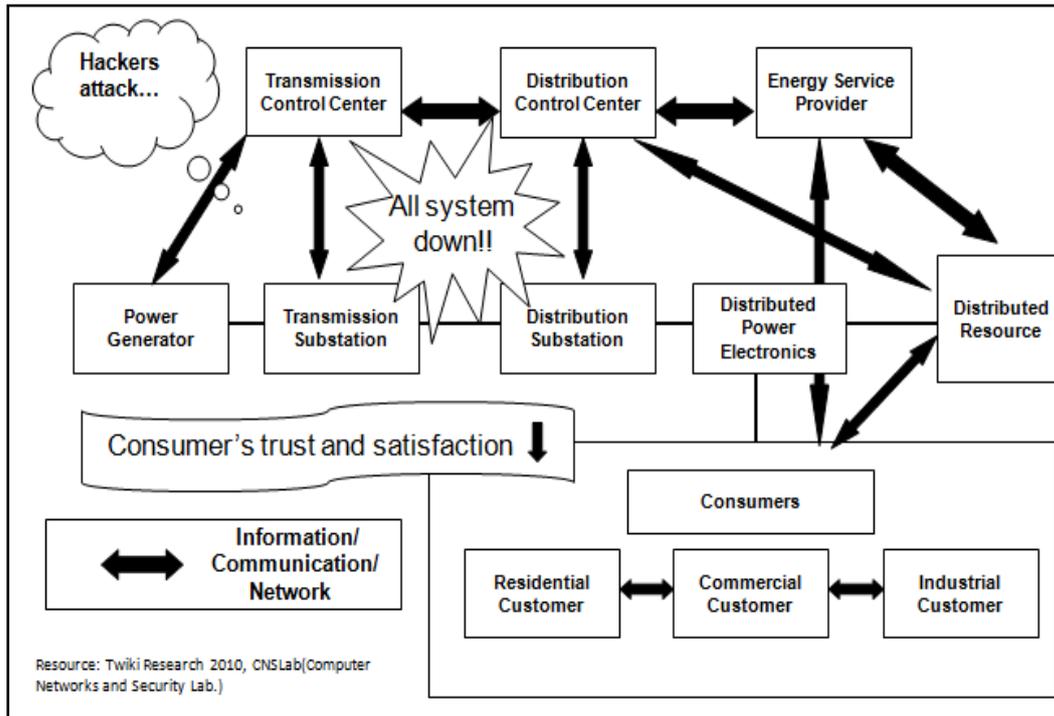


Figure. 1. Security Importance [5].

As the consumer’s awareness and participation in the smart grid project has increased, it is interesting to note what criteria held by the consumer will impact the information security system. The importance of information security criteria was the main aspect perceived to impact customer trust towards the entire smart grid system [6], as shown in figure 1.

2.3 Smart Grid Conceptual Mind Mapping

Figure 2 below showed Smart grid conceptual mapping where the study area was highlighted. Under the umbrella of Smart Planet cultivated by IBM, lay Smart Water Management and Green Planet, Smarter Planet Skill and Education, Smart Grid, Smart Health Care and Smart Cities. There were eight main elements in Smart Grid, narrowing down to network security utilities, business network, communication and information security’s impact on consumer trust and satisfaction which was the main focus of this paper. Smart grid utilizes communication technology and information to optimally transmit and distribute electricity from suppliers to consumers.

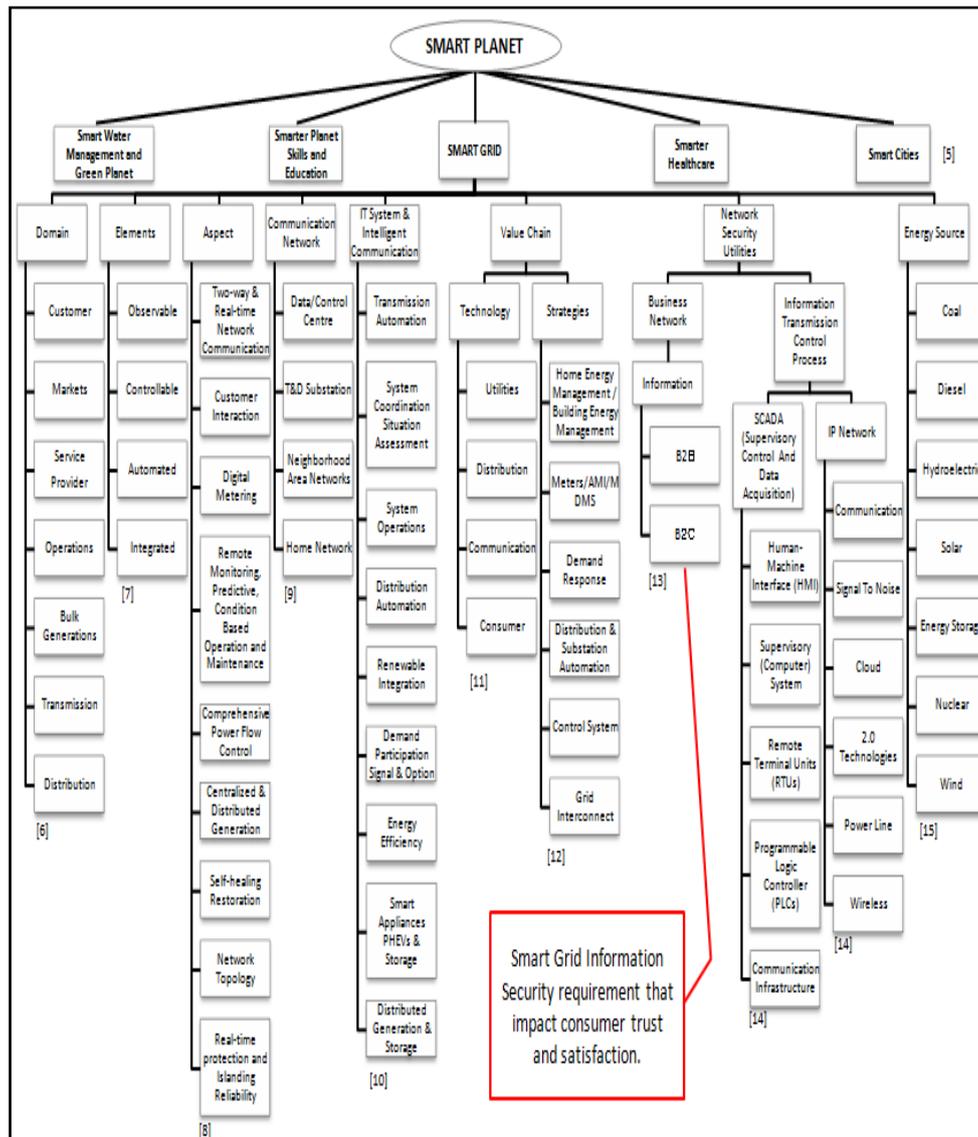


Figure. 2. Smart Grid Conceptual Mind Mapping [6].

2.4 The need to protect the privacy and security of priceless data

Personal information leakage incidents could be vital and cost a big sum of compensation to the utilities. The number of victims of information leakage incidents for 2008 amounted to 7.23 million individuals, the following table 1 shows the summary of data collected for 2008 in the Information Security Incident Survey Report of NPO Japan Network Security Association Security Incident Investigation Working Group.

Table 1. Summary Data of 2008 Personal Information Leakage Incidents [7]

Number of Victims	7,232,763
Number of Incidents	1,373
Total Projected Compensation for Damages	¥236,725,290,000
Number of Victims per Incident^{*1}	5,668
Average Projected Compensation for Damages per Incident^{*1}	¥185,520,000
Average Projected Compensation for Damages per Victim^{*2}	¥43,632

A number of smart grid information security requirements and regulations were available online, those guidelines were a significant step in securing the smart grid but they do not fully address potential vulnerabilities that can emerge [8, 9]. The strongest arguments made for securing smart meters, saying that consumers will have physical and potentially logical access to the smart meters [10]. Security was generally described in terms of availability, integrity, and confidentiality. Cyber systems were observed to be vulnerable to worms, viruses, denial-of-service attacks, malware, phishing, and user errors that compromise integrity and availability [11].

It is important to avoid bearing any potential security risk. As a matter of fact, grid security technologies have been so far designed and deployed as a middleware layer add-on, independently at each tier [12]. The need to protect the privacy and security of priceless data over the grid was fueling even more need for common security evaluation criteria. In brief, information security professionals need to be aware that the workings of the most basic IT resource of electricity supply is changing in a manner that introduces a far larger and remotely addressable attack surface combined with the tempting opportunity for mischief and monetary gain [13].

III. 3 SIGNIFICANCE OF THE STUDY

Because a smart grid utilizes digital technology to provide two-way communication between suppliers and consumers home electronics through the use of smart meters, information security protection measures need to be consumer friendly and practicable to be implemented on all level within a community or organization [14]. For this reason, it is important for us to examine on the voice of customer on their agreement on the essentiality of the identified sixteen consumer requirements [6], and the ranking of essentiality of the requirements. The correlation relationship between the requirements will show how a trend will change when a requirement is being implemented or abolished, aim is to provide a deeper investigation into the information cues on security requirements that really attract consumer interest and therefore may merit specific attention in future information security policies.

IV. METHODOLOGY

4.1 Hermeneutic Circle: Consumer Requirements Identified

Hermeneutic Circle is an interpretive and conceptual-analytical research method. Hermeneutics refers to the idea that one understands of the text as a whole is established by reference to the individual parts and one understands of each individual part in reference to the whole [15, 16, 17]. The information security consumer requirements were carefully picked from literature review, referring to experts' opinions and focused group discussion.

Table 2 referred to the sixteen requirements that were identified via the work of voice of customer and careful literature review based on Hermeneutic Circle Methodology classed in five categories of philosophy, human behavior, rule based social system, strategic system and hardware.

Table 2. Information Security Consumer Requirements and its Significant

Variable	Category	IS Consumer Requirement	Significance
1	Philosophy	Confidentiality	Unauthorized disclosure of information.
2		Integrity	Unauthorized modification or destruction of information; Strong requirement that information should not be modified by unauthorized entities, and should be validated for accuracy and errors.
3		Availability	Disruption of access to or use of information or an information system; Strong requirement that information should be available within appropriate time frames.
4		Privacy concern	Strong requirement that information should not be viewed by unauthorized entities.
5	Human Behavior	Tactical oversight monitoring system	Monitoring and reporting of the security status or posture of an IT system can be carried out to determine compliance with security requirements (e.g., policy, procedures, and regulations), gauge the effectiveness of security controls and manage risk.
6		Facilities misuse prevention	Organization monitored for unauthorized use of information processing facilities, users aware of their exact scope of permitted access. Users aware that monitoring tools are being used to detect unauthorized use.
7	Rule base Social System	Networking issues	Investigate ways to ensure that commercially available components, public networks like the Internet, or available enterprise systems can be implemented without jeopardizing security or reliability.
8		Quality assurance	Security metrics can be used during the software development lifecycle to eliminate vulnerabilities, particularly during code production, by performing functions such as measuring adherence to secure coding standards, identifying likely vulnerabilities.
9		Mature or proprietary protocols	Immature or proprietary protocols may not be adequately tested either against inadvertent compromises or deliberate attacks. This may leave the interface with more vulnerabilities than if a more mature protocol were used.
10	Strategic System	Cryptography and key management	To enable key management on a scale involving, potentially, tens of millions of credentials and keys as well as local cryptographic processing on the sensors such as encryption and digital signatures.
11		Reliable systems level	Where research on a number of related topics is required to further approaches to building advanced protection architecture that can evolve and can tolerate failures, perhaps of a significant subset of constituents.
12		Strategic support	Assessments of security properties can be used to aid different kinds of decision making, such as program planning, resource allocation, and product and service selection.
13		Security in wireless media	Wireless media may necessitate specific types of security technologies to address wireless vulnerabilities across the wireless path.
14	Hardware	Reliable device level	Efforts to devise cost-effective, tamper-resistant architectures for smart meters and other components, which are necessary for systems-level survivability and resiliency and for improving intrusion detection in embedded systems.
15		High bandwidth of communications channels	Severely-limited bandwidth may constrain the types of security technologies that should be used across an interface while still meeting that interface's performance requirements.
16		Microprocessor perform memory and compute capabilities	Severely-limited memory and/or compute capabilities of a microprocessor-based platform may constrain the types of security technologies, such as cryptography, that may be used while still allowing the platform to meet its performance requirements.

4.2 Questionnaire Construction

Preferences questionnaires with questions that measures separate variables were adopted to generate the bounded questionnaire where the respondent was presented with a continuous likert scale. A non-comparative likert scaling

techniques was used. The level of measurement of a variable in mathematics and statistics was a classification that was proposed in order to describe the nature of information contained within numbers assigned to objects and, therefore, within the variable.

The questionnaire was divided into two sections:

1. Consumer's perception on Consumer Requirements Model
2. Demography

The respondent was asked to indicate his or her degree of agreement with the statement or any kind of subjective or objective evaluation of the statement. In Section 1, a five-point likert scale was used, rating value from strongly agreed, agreed, neutral, disagreed, and strongly disagreed. The questions comprised sixteen elements such as confidentiality, integrity, availability, privacy concern, tactical oversight monitoring system, facilities misuse prevention, networking issues, quality assurance, mature or proprietary protocols, cryptography and key management, reliable system level, strategy support, security in wireless media reliable device level, high bandwidth of communication channels, and microprocessor perform memory and compute capabilities. The demography variables measured at a nominal level in Section 2 included information on area of research interest and countries.

4.3 Data Distribution and Collection

50 sets of questionnaire were distributed via international workshop, email and peer groups. The selection of the respondents was carefully done based on respondents who possesses limited or more knowledge on the subject of smart grid and its information security. The unreturned questionnaire recorded 7sets; returned usable questionnaire recorded 38; and the returned un-usable questionnaire recorded 5. Figure 3 showed the work flow of questionnaire development, distribution, collection and analysis. Each of the transaction numbers were recorded for back-up references. The returned usable 38 set questionnaires were used to generate analysis.

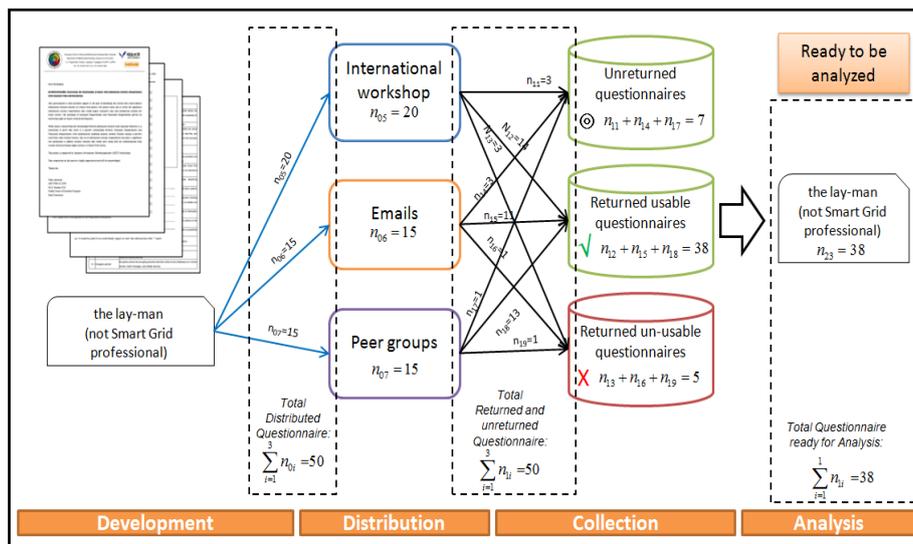


Figure. 3. Questionnaire Development, Distribution, Collection and Analysis.

4 Data Analysis

4.4.1 Descriptive Statistic and Ranking of Variable

Descriptive statistics was adopted as the main methodology for analysis [18]. Correlation was used to analyze statistical relationships between the consumer requirements' variables with the data provided from the questionnaires collected to indicate a predictive relationship that can be exploited in practice.

Minitab was employed to enhance the mathematical calculation to validate the effects of the importance of the consumer of the information security in a smart grid system. Variables of requirements were evaluated one by one so that a ranking has to be chosen by considering only the mean value from the result of descriptive analysis of the raw data obtained from the returned questionnaire.

4.4.2 Histogram and Individual Value Plot versus Country

Histogram is useful in visualizing the consensus of the respondents in this paper such as distribution and data trend. We proposed to adopt histogram method for the decision making to support the hypothesis [19]. Individual value plot was employed to set the appropriate course for the analysis and displayed all data values for small amounts of data. The data collected from different country was been analyzed on the respondents perception on their agreement on the selected consumer requirements.

4.4.3 Chi Square Test (χ^2 Test)

χ^2 Test was employed as the statistical hypothesis test for the reason of our sampling distribution of the test was a chi-square distribution and the sampling distribution can be made to approximate a chi-square distribution as closely as desired by making the sample size large enough [20].

4.4.4 Pearson Correlation and Degree of Correlation

To obtain the optimal result with respect to the Pearson correlation, we use the Minitab software to generate its output value with the analysis of the respective 16 requirements. An assumption of value lower than 0.5 has low correlation was necessary to analyze and find out which variable is not correlated to one another or having a low correlation relationship to other variable. The Pearson correlation data was used to evaluate the correlation degree and it was defined as strong-normal-weak where S: Strong (≥ 0.6), N: Normal (0.4-0.6) and W: Weak (≤ 0.4).

4.4.5 Lexicographical Order

The output value of Pearson correlation was employed to rank the variables using the lexicographical order methodology. An important property of the lexicographical order in this paper is that it preserves well-orders, when applied to permutations, lexicographic order is increasing numerical order [21].

4.4.6 Network Relationship Ring, Cluster and Tree

Correlation strength was evaluated through stimulation with a wide range of raw data, by varying the network size, number of variables, and the mean values using the Cytoscape software. The network relationship ring achieved two objectives: Firstly to minimize the maximum nodal degree and eliminate low correlated variables, and secondly to minimize the total degree over all network nodes to find strong correlation between variables. Then the network was divided into two clusters and a tree to demonstrate the correlation relationship of the variables perceived by consumers.

V. RESULTS

5.1 Descriptive Statistics

This part of analysis quantitatively described the main features of the collected data and provided a snap shot of the situation under this study. The descriptive statistics chosen include: N, N*, Percent, CumPct, Mean, SE Mean, Standard Deviation (StDev), Sum, Minimum, Q1, Median, Q3, Maximum, and Range. N was used to track down if any cases were being lost between variables.

Focused on the mean showed in table 3, variable four had the highest mean of 4.7632, which means most of the respondents favor and perceived that variable four had the essentiality as a requirement for a sustainable information security system. On the other hand, tracing variable fifteen, it had the lowest mean of 4.079; this could be one of the less important requirement contributes to the sustainable of the information security in a smart rid system. Where both variable having the same mean, standard deviation (StDev) was taken into account where lower value of StDev rank higher.

Table 3. Descriptive Statistic Table

Descriptive Statistics: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16															
Ranking	Variable	N	N*	Percent	CumPct	Mean	SE Mean	StDev	Sum	Minimum	Q1	Median	Q3	Maximum	Range
2	1	38	0	100	100	4.737	0.105	0.644	180.000	2.000	5.000	5.000	5.000	5.000	3.000
3	2	38	0	100	100	4.605	0.110	0.679	175.000	2.000	4.000	5.000	5.000	5.000	3.000
6	3	38	0	100	100	4.500	0.140	0.862	171.000	1.000	4.000	5.000	5.000	5.000	4.000
1	4	38	0	100	100	max 4.7632	0.0794	0.4896	181.0000	3.0000	5.0000	5.0000	5.0000	5.0000	2.0000
12	5	38	0	100	100	4.132	0.108	0.665	157.000	3.000	4.000	4.000	5.000	5.000	2.000
4	6	38	0	100	100	4.526	0.111	0.687	172.000	3.000	4.000	5.000	5.000	5.000	2.000
10	7	38	0	100	100	4.289	0.113	0.694	163.000	2.000	4.000	4.000	5.000	5.000	3.000
7	8	38	0	100	100	4.474	0.118	0.725	170.000	2.000	4.000	5.000	5.000	5.000	3.000
15	9	38	0	100	100	4.079	0.127	0.784	155.000	2.000	4.000	4.000	5.000	5.000	3.000
13	10	38	0	100	100	4.132	0.126	0.777	157.000	2.000	4.000	4.000	5.000	5.000	3.000
8	11	38	0	100	100	4.474	0.124	0.762	170.000	2.000	4.000	5.000	5.000	5.000	3.000
11	12	38	0	100	100	4.184	0.112	0.692	159.000	3.000	4.000	4.000	5.000	5.000	2.000
5	13	38	0	100	100	4.526	0.124	0.762	172.000	2.000	4.000	5.000	5.000	5.000	3.000
9	14	38	0	100	100	4.368	0.109	0.675	166.000	2.000	4.000	4.000	5.000	5.000	3.000
16	15	38	0	100	100	min 4.079	0.157	0.969	155.000	2.000	3.000	4.000	5.000	5.000	3.000
14	16	38	0	100	100	4.132	0.137	0.844	157.000	2.000	4.000	4.000	5.000	5.000	3.000

Terms in the output and some definitions

N = number of data items in the sample

N* = number of items in the sample that have missing values

Mean = average

Median = 50th percentile

TrMean= the 5% trimmed mean

StDev = standard deviation

SE Mean = standard error of the mean = standard deviation divided by the square root of the sample size

Minimum = smallest data value

Maximum = largest data value

Q1 = 25th percentile = first quartile

Q3 = 75th percentile = third quartile

5.2 Ranking of Variable: Information Security Consumer Requirement

The consumer requirements from table 2 were numbered after the variables' rank referred to the result from Descriptive Statistics analysis, as showed in table 4.

Table 4. Information Security Consumer Requirement

Variable	Category	IS Consumer Requirements	Degree of Ranking
1	Philosophy	Confidentiality	2
2		Integrity	3
3		Availability	6
4		Privacy concern	1
5	Human Behavior	Tactical oversight monitoring system	12
6		Facilities misuse prevention	4
7	Rule base Social System	Networking issues	10
8		Quality assurance	7
9		Mature or proprietary protocols	15
10	Strategic System	Cryptography and key management	13
11		Reliable systems level	8
12		Strategic support	11
13		Security in wireless media	5
14	Hardware	Reliable device level	9
15		High bandwidth of communications channels	16
16		Microprocessor perform memory and compute capabilities	14

Then all the requirements were then again rearranged according to the ranked number, as in table 5. Now, the essentiality ranking of the consumer requirements was clearly displayed in a column. Privacy concern apparently has the potential to become the most important element in among the other fifteen requirements in terms of information security in a smart grid system and high bandwidth of communications channels ranked last.

Table 5. Essentiality Ranking of Information Security Consumer Requirement

Essentiality Ranking	Consumer Requirements
1	Privacy concern
2	Confidentiality
3	Integrity
4	Facilities misuse prevention
5	Security in wireless media
6	Availability
7	Quality assurance
8	Reliable systems level
9	Reliable device level
10	Networking issues
11	Strategic support
12	Tactical oversight monitoring system
13	Cryptography and key management
14	Microprocessor perform memory and compute capabilities
15	Mature or proprietary protocols
16	High bandwidth of communications channels

Comparison was made between table 4 and table 5, the result showed that each of the ranking was a mixture of from five categories of philosophy, human behavior, rule based social system, strategic system and hardware, and hence each category could be concluded to have a balance value of essentiality.

5.3 Histogram

Figure 4 graphically represented a visual impression of the distribution of data and an estimate of the probability distribution of a continuous variable, showed the distribution of a quantitative variable by its relative frequency of data points in an interval. The result output presented a bell-curve slanting to the right, fitting the normal distribution. There was no strong deviation from the distribution, and there was adequate data cleaning. This indicated that the average respondents was at least ticking the value of more than average, that was on agreed or strongly agreed side when asked about the importance of identified requirement in conjunction to the information security in a smart grid system.

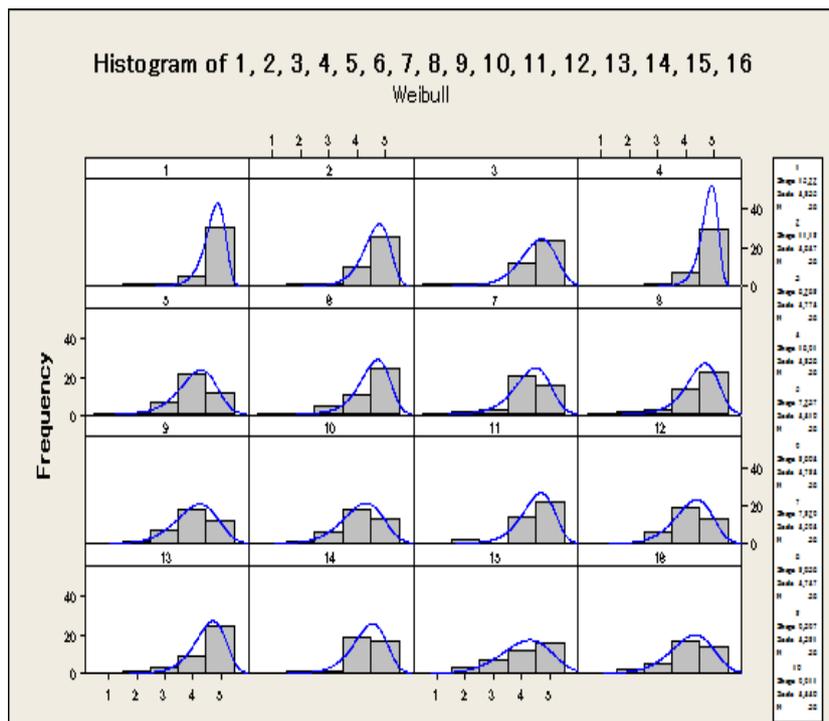


Figure 4. Histogram.

5.4 Chi-Square Test

To prove that all variables are slanted to the right, we made hypothesis as per below and tested it with Chi-Square test. Hypothesis was stated as per below:

Null Hypothesis

H_0 = All 16 graph of variables is as expected value

*Explanation = All identified information security consumer requirements are important

Alternative Hypothesis

H_1 = Not all 16 graph of variables slant to the right

*Explanation = Not all identified information security consumer requirements are important

Functions are defined as follow:

f_α : Total number from the sample which chosen answer 1, 2 and 3.

f_β : Total number from the sample which chosen answer 4 and 5.

p_α : Expected percentage of the answer 1, 2 and 3.

p_β : Expected percentage of the answer 4 and 5.

The philosophy and strategy of management with regard to information security is the perfect standard against which technology and other security mechanisms can be measured [22]. Due to this reason, we considered the null hypothesis for philosophy category as $p_\beta = 0.95$ slightly higher than the rest category as $p_\beta = 0.85$.

5.4.1 Chi-square test for the class of “Philosophy” (Variable 1-4)

For the class of “Philosophy”, we considered the null hypothesis as following

$$H_0 : p_\alpha = 0.05, \quad p_\beta = 0.95 .$$

The formula of Chi-square test [20] is

$$\chi^2 = \frac{(f_\alpha - n * p_\alpha)^2}{n * p_\alpha} + \frac{(f_\beta - n * p_\beta)^2}{n * p_\beta}$$

Table 6. Chi-Square Test for the Class of “Philosophy” (Variable 1)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
2	36	1.9	36.1	0.0055

Chi-square value $\chi^2 = 0.0055 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 7. Chi-Square Test for the Class of “Philosophy” (Variable 2)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
2	36	1.9	36.1	0.0055

Chi-square value $\chi^2 = 0.0055 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 8. Chi-Square Test for the Class of “Philosophy” (Variable 3)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
2	36	1.9	36.1	0.0055

Chi-square value $\chi^2 = 0.0055 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 9. Chi-Square Test for the Class of “Philosophy” (Variable 4)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
1	37	1.9	36.1	0.4488

Chi-square value $\chi^2 = 0.4488 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

5.4.2 Chi-square test for the class of “Human behavior” (Variable 5-6)

For the class of “Human behavior”, we considered the null hypothesis as following

$$H_0 : p_\alpha = 0.15, \quad p_\beta = 0.85 .$$

The formula of Chi-square test is

$$\chi^2 = \frac{(f_\alpha - n * p_\alpha)^2}{n * p_\alpha} + \frac{(f_\beta - n * p_\beta)^2}{n * p_\beta}$$

Table 10. Chi-Square Test for the Class of “Human Behaviour” (Variable 5)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
5	33	5.7	32.3	0.1011

Chi-square value $\chi^2=0.1011 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 11. Chi-Square Test for the Class of “Human Behaviour” (Variable 6)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
4	34	5.7	32.3	0.5965

Chi-square value $\chi^2=0.5965 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

5.4.3 Chi-square test for the class of “Rule based on social system” (Variable 7-9)

For the class of “Rule based on social system”, we considered the null hypothesis as following

$$H_0 : p_\alpha = 0.15, \quad p_\beta = 0.85 .$$

The formula of Chi-square test is

$$\chi^2 = \frac{(f_\alpha - n * p_\alpha)^2}{n * p_\alpha} + \frac{(f_\beta - n * p_\beta)^2}{n * p_\beta}$$

Table 12. Chi-Square Test for the Class of “Rule Based on Social System” (Variable 7)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
3	35	5.7	32.3	1.5046

Chi-square value $\chi^2=1.5046 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 13. Chi-Square Test for the Class of “Rule Based on Social System” (Variable 8)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
3	35	5.7	32.3	1.5046

Chi-square value $\chi^2=1.5046 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 14. Chi-Square Test for the Class of “Rule Based on Social System” (Variable 9)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
8	30	5.7	32.3	1.0918

Chi-square value $\chi^2=1.0918 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

5.4.4 Chi-square test for the class of “Strategic system” (Variable 10-13)

For the class of “Strategic system”, we considered the null hypothesis as following

$$H_0 : p_\alpha = 0.15, \quad p_\beta = 0.85 .$$

The formula of Chi-square test is

$$\chi^2 = \frac{(f_\alpha - n * p_\alpha)^2}{n * p_\alpha} + \frac{(f_\beta - n * p_\beta)^2}{n * p_\beta}$$

Table 15. Chi-Square Test for the Class of “Strategic System” (Variable 10)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
7	31	5.7	32.3	0.3488

Chi-square value $\chi^2=0.3488 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 16. Chi-Square Test for the Class of “Strategic System” (Variable 11)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
3	36	5.7	32.3	2.8256

Chi-square value $\chi^2=2.8256 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 17. Chi-Square Test for the Class of “Strategic System” (Variable 12)

f_α	f_β	$n * p_\alpha$	$n * p_\beta$	χ^2
6	32	5.7	32.3	1.0186

Chi-square value $\chi^2=0.0186 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 18. Chi-Square Test for the Class of “Strategic System” (Variable 13)

f_{α}	f_{β}	$n * p_{\alpha}$	$n * p_{\beta}$	χ^2
4	34	5.7	32.3	0.5965

Chi-square value $\chi^2=0.5965 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

5.4.5 Chi-square test for the class of “Hardware” (Variable 14-16)

For the class of “Hardware”, we considered the null hypothesis as following

$$H_0 : p_{\alpha} = 0.15, \quad p_{\beta} = 0.85 .$$

The formula of Chi-square test is

$$\chi^2 = \frac{(f_{\alpha} - n * p_{\alpha})^2}{n * p_{\alpha}} + \frac{(f_{\beta} - n * p_{\beta})^2}{n * p_{\beta}}$$

Table 19. Chi-Square Test for the Class of “Hardware” (Variable 14)

f_{α}	f_{β}	$n * p_{\alpha}$	$n * p_{\beta}$	χ^2
2	36	5.7	32.3	2.8256

Chi-square value $\chi^2=2.8256 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 20. Chi-Square Test for the Class of “Hardware” (Variable 15)

f_{α}	f_{β}	$n * p_{\alpha}$	$n * p_{\beta}$	χ^2
10	28	5.7	32.3	3.8163

Chi-square value $\chi^2=3.8163 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

Table 21. Chi-Square Test for the Class of “Hardware” (Variable 16)

f_{α}	f_{β}	$n * p_{\alpha}$	$n * p_{\beta}$	χ^2
7	31	5.7	32.3	0.3488

Chi-square value $\chi^2=0.3488 < 3.841 = \chi^2_{0.05,1}$, the hypothesis H_0 was not rejected.

From the result above, we concluded that the entire null hypothesis from each class has not been rejected. Also, we obtained the following result:

$$p_{\beta} > 0.8$$

This means that results of p_{β} showed that the value is greater than 80% of the entire respondents’ agreement on the essentiality of the consumer requirements identified. Therefore, this result supported our null hypothesis that the graphs of all of the variables are slanted to the right.

In conclusion, the null hypothesis was not rejected.

5.5 Individual Value Plot versus Country

There were respondents from 14 countries took part in the questionnaire activity. Individual value plot versus country were accomplished in purpose to identify the difference of perception of consumers from different countries, as showed in figure 5.

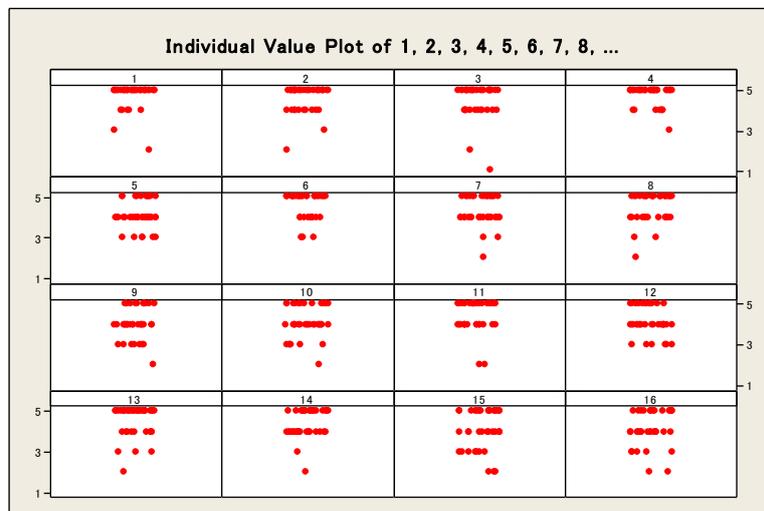


Figure. 5. Summary of Individual Value Plot versus Countries.

5.5.1 Confidentiality versus Country

Figure 6 showed most of the countries strongly agreed that confidentiality was essential for a sustainable information security system in a smart grid; there was a dot on value 2 for the country of China. This could be due to the confidentiality in China has been a volatile for the country.

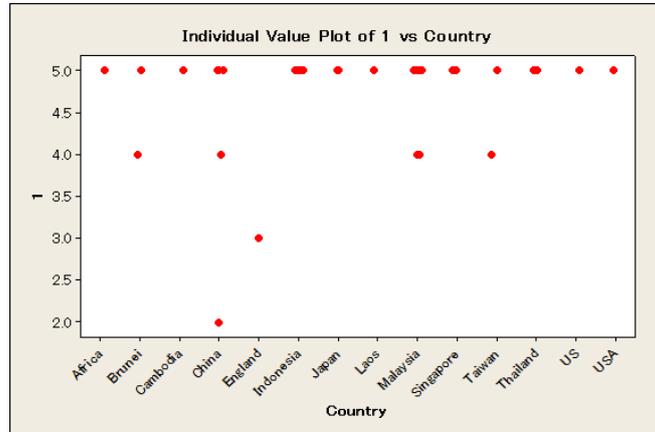


Figure. 6. Individual Value Plot of Confidentiality versus Countries.

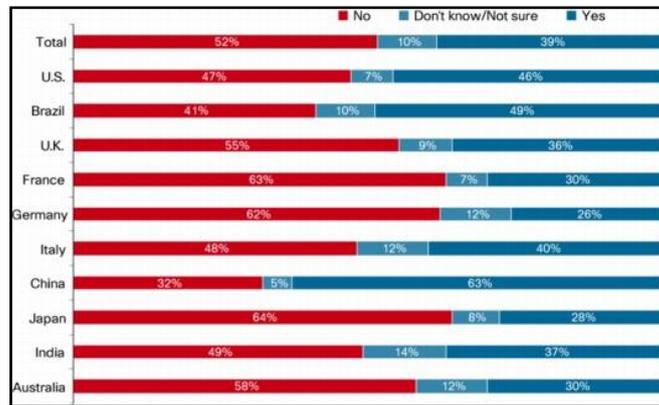


Figure 7 showed the number of times IT has had to deal with an employee for accessing unauthorized networks or facilities:

Figure. 7. Employee Accessed Unauthorized Networks or Facilities.

39 percent of IT professionals said they have dealt with an employee accessing unauthorized parts of a company's network or facility, with almost half of IT professionals reporting this in Brazil (49 percent) and the United States (46 percent), and 63 percent in China [23]. This could support on the dot on value 2 for China being stating that confidentiality was not so important contribute towards the sustainability of an information security in a smart grid system.

5.5.2 Reliable systems level versus Country

Among the thirty eight respondents, only two, from China and Taiwan, disagreed where research on a number of related topics was required to further approaches to building advanced protection architecture that can evolve and can tolerate failures, as showed in figure 8. This might due to the rapidly growing economies like China and Taiwan have different smart grid infrastructure needs from those of OECD countries [24].

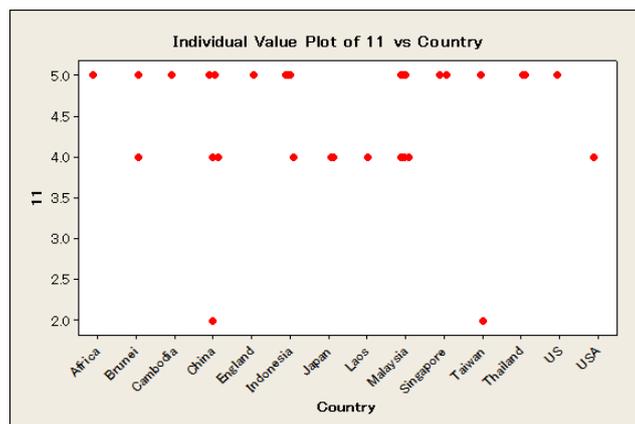


Figure. 8. Value Plot of Reliable systems level versus Countries.

5.5.3 Microprocessor performs memory and compute capabilities versus Country

There were relatively 5 respondents choose to be in neutral and two from Malaysia and Taiwan were disagreed that microprocessor performs memory and compute capabilities was an essential element for an information security in a smart grid system, as showed in figure 9.

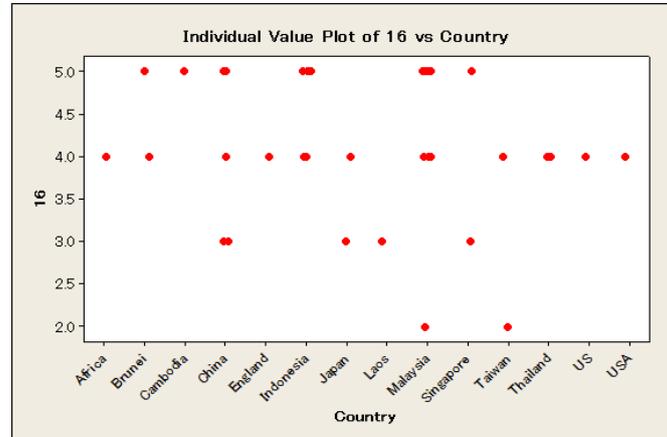


Figure. 9. Value Plot of Microprocessor performs memory and compute capabilities versus Countries.

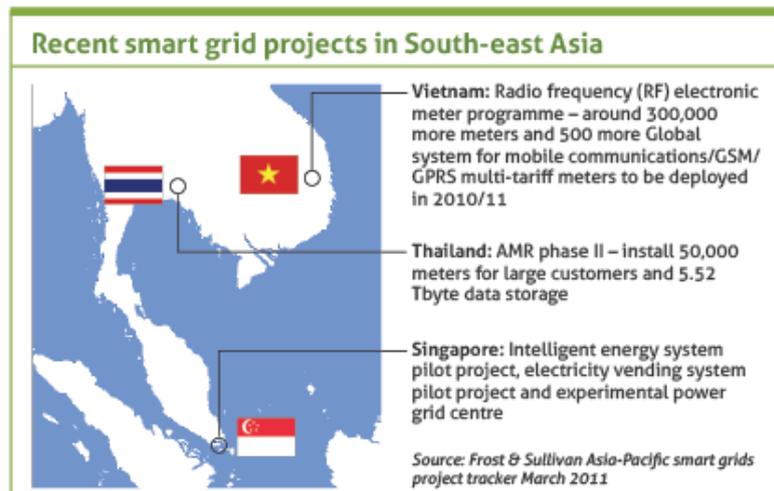


Figure. 10. Recent Smart Grid Projects in South-east Asia.

Malaysia is a developing country and was being lobbied recently to develop a roadmap and institutional framework to ensure coordinated efforts for the long-term. Figure 10 showed the recent smart grid projects in South-east Asia [25]. The respondent might not aware of the importance of what the microprocessor performs memory and compute capabilities could bring effect to the system. As for Taiwan, it is the home to the world's largest contract manufacturing firms, accounting for 65% of the world's Electronics Manufacturing Services business [26]. Taiwan has already made the semiconductor industry akin to the contract manufacturing industry, thus the quality of microprocessor is assured.

In conclusion, the result of individual value plot versus country shows that there are still minority respondents stating that confidentiality is not so important contribute towards the sustainability of an information security in a smart grid system, disagreed where research on a number of related topics was required to further approaches to building advanced protection architecture that can evolve and can tolerate failures, and disagreed that microprocessor performs memory and compute capabilities was an essential element for an information security in a smart grid system. This might due to the rapidly growing economies in different countries have different smart grid infrastructure needs.

5.6 Pearson Correlation

In this section, we examined the relationship between two variables. The Pearson Correlation was employed in the computation process. Result in figure 11 showed that most of the variable of the requirements were positive correlated to each other. Variable 13, security in wireless media, seems to have no strong correlation with other fifteen variables. This could be due to its entertainment element contributed less relationship with other information security requirement although the essentiality of security in wireless media was essential.

There was some reason supported this findings. Wireless networks offer great potential for exploitation for two reasons; they use the airwaves for communication, and wireless-enabled laptops are ubiquitous. Several security protocols designed for wired-line networks have been adopted for use in wireless networks. However, they may not be suitable for wireless

networks and devices since scenarios and capabilities applicable to wired-line networks may not be valid in wireless networks [27]. Most security technologies are currently deployed in wired networks and are not fully applicable to wireless networks involving mobile devices with limited computing capability [28].

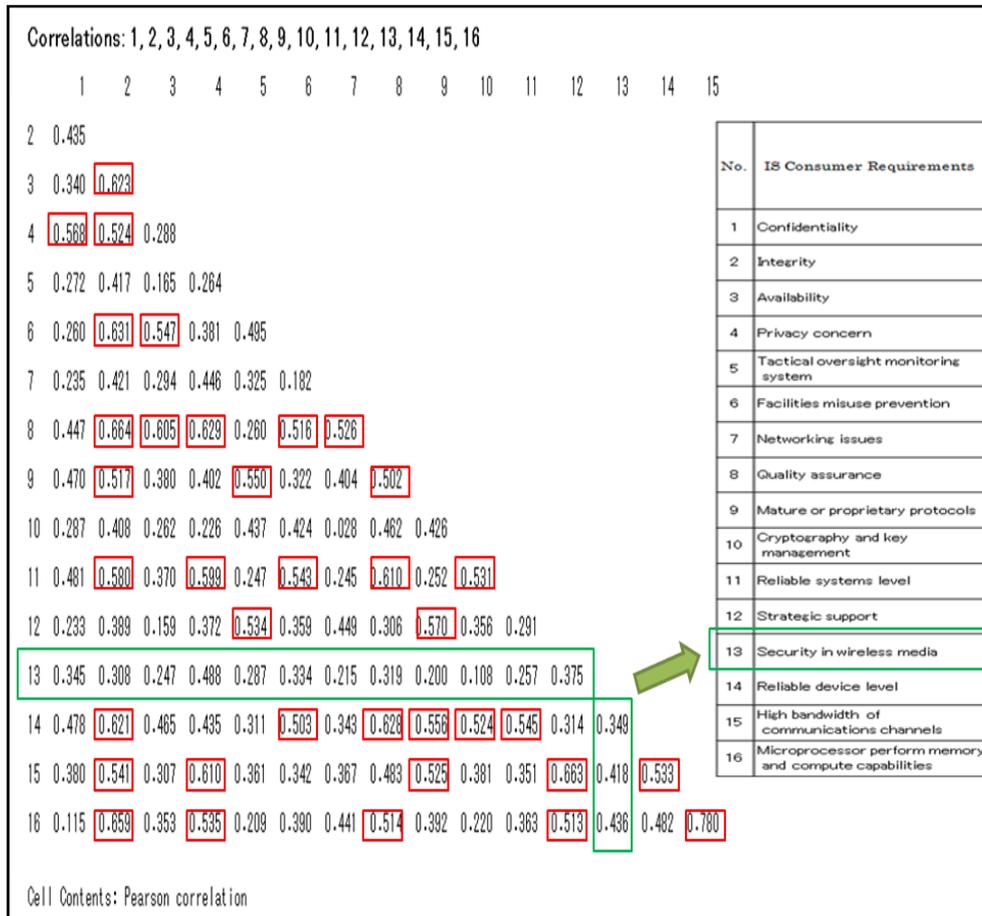


Figure 11. Pearson Correlation Test with P-value test.

In conclusion, we could see that the wireless connections have become an increasingly popular way of providing internet access although it is still in its development process. When the wireless becomes a common method for data transmission especially from the smart meter to the grid, security becomes a highly important part of the wireless network structure. Hence, although the security in wireless media is not correlated to other variable tested, it is still an important asset for all systems using a wireless network.

5.6.1 Degree of Correlation

The degree of correlation analysis was worked out to consider how strong or how weak each pair of tested variables were correlated among each other, as showed in table 22. We fixed three categories explicating strong, normal and weak for each pair of tested variables, as per below.

- S: Strong (≥ 0.6)
- N: Normal (0.4-0.6)
- W: Weak (≤ 0.4).

Table 22. Degree of Correlation Analysis.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		N	W	N	W	W	W	N	N	W	N	W	W	N	W	W
2	N		S	N	N	S	N	S	N	N	N	W	W	S	N	S
3	W	S		W	W	N	W	S	W	W	W	W	W	N	W	W
4	N	N	W		W	W	N	S	N	W	N	W	N	N	S	N
5	W	N	W	W		N	W	W	N	N	W	N	W	W	W	W

6	W	S	N	W	N		W	N	W	N	N	W	W	N	W	W
7	W	N	W	N	W	W		N	N	W	W	N	W	W	W	N
8	N	S	S	S	W	N	N		N	N	S	W	W	S	N	N
9	N	N	W	N	N	W	N	N		N	W	N	W	N	N	W
10	W	N	W	W	N	N	W	N	N		N	W	W	N	W	W
11	N	N	W	N	W	N	W	S	W	N		W	W	N	W	W
12	W	W	W	W	N	W	N	W	N	W	W		W	W	S	N
13	W	W	W	N	W	W	W	W	W	W	W	W		W	N	N
14	N	S	N	N	W	N	W	S	N	N	N	W	W		N	N
15	W	N	W	S	W	W	W	N	N	W	W	S	N	N		S
16	W	S	W	N	W	W	N	N	W	W	W	N	N	N	S	

The accumulated numbers of S, N and W were calculated for each variable and were recorded in table 23.

Table 23. Accumulated Numbers of the Degree of Correlation Analysis.

	S	N	W
1	0	6	9
2	5	8	2
3	2	2	11
4	2	8	5
5	0	5	10
6	1	6	8
7	0	6	9
8	5	7	3
9	0	10	5
10	0	7	8
11	1	6	8
12	1	4	10
13	0	3	12
14	2	9	4
15	3	5	7
16	2	6	7

5.6.2 Lexicographical Order

Upon the completion of table 23, correlation ranking of information security consumer requirements via pearson correlation test observation was generated employing lexicographical order methodology.

Table 23 showed that variable 2 was larger than 8, followed by 15, 14, 4, 16, 3, 11, 6, 12, 9, 10, 1, 7, 5, 13. In the process of calculation, the numbers in S column will be taken to identify which was the largest numbers among the variable which gain most S means it has the strongest correlation value to other pairs of variable being compared. When it comes to same number counted, for example, variable 2 and variable 5 has the same S count of five, and then second column of N will be referred, for this case, N counts of variable 2 was larger than variable 8, so variable 2 was considered to be larger than variable 8.

Lexicographical Ordering was adopted to create result as per below.

Variable 2 > 8 > 15 > 14 > 4 > 16 > 3 > 11 > 6 > 12 > 9 > 10 > 1 > 7 > 5 > 13

A table of correlation ranking of information security consumer requirements via Pearson correlation test observation was developed, as showed in table 24.

Table 24. Table of Correlation Ranking of Information Security Consumer Requirements.

Correlation Ranking	Consumer Requirements
1	Integrity
2	Quality assurance
3	High bandwidth of communications channels
4	Reliable device level
5	Privacy concern
6	Microprocessor perform memory and compute capabilities
7	Availability
8	Reliable systems level
9	Facilities misuse prevention
10	Strategic support
11	Mature or proprietary protocols
12	Cryptography and key management
13	Confidentiality
14	Networking issues
15	Tactical oversight monitoring system
16	Security in wireless media

Table 24 concluded that integrity has the strongest correlation strength with other pair of information security consumer requirement identified, followed by quality assurance, high bandwidth of communication channels, reliable device level, privacy concern, microprocessor perform memory and compute capabilities, availability, reliable system level, facilities misuse prevention, strategy support, mature or proprietary protocols, cryptography and key management, confidentiality, networking issues, tactical oversight monitoring system, and last but not least, security in wireless media.

The purpose of this ranking is to create a picture of a common language and a detailed understanding of the customer's requirements. When a greater importance is placed on consumer retention, the system can be developed to gain their expectations, for an enlightened consumer is an empowered consumer. Instead of adding to the previous debate whether and how many consumers are interested in information security as such, the aim is to provide a deeper investigation into the information cues on security requirements that really attract consumer interest and therefore may merit specific attention in future information security policies. Statistic sometimes shows differences from the theoretical result that we might not know. Based on this study, the result showed that integrity, quality assurance and high bandwidth of communications channels has high correlation with each other. We could conclude that any changes in the unauthorized modification or destruction of information will lead to the changes in the quality assurance performing functions and types of security technologies that should be used across an interface while still meeting that interface's performance requirements.

5.6.3 Correlation Relationship Ring

We study the dynamics of correlation between the variables with different value of threshold. For many such situations, it was found that the correlations between individual variables are better indicators than the value of attributes.

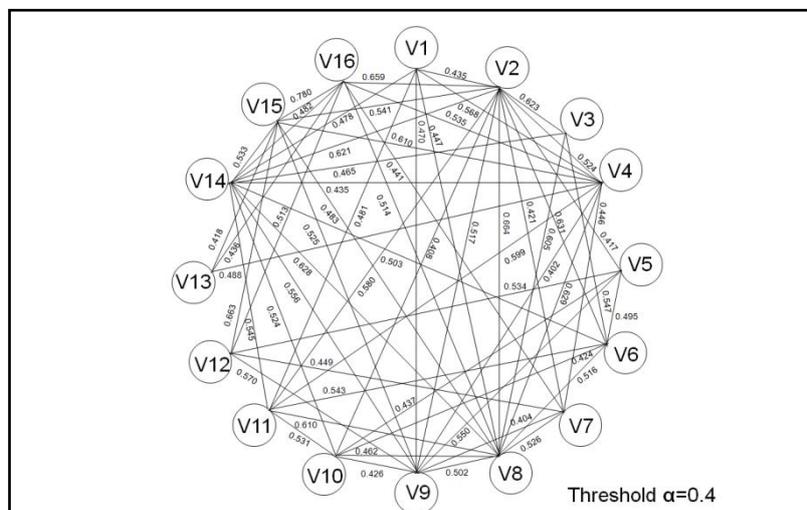


Figure. 12. Correlation Relationship Ring, Threshold $\alpha = 0.4$.

Figure 12 above shows the correlation with the threshold α value of 4.

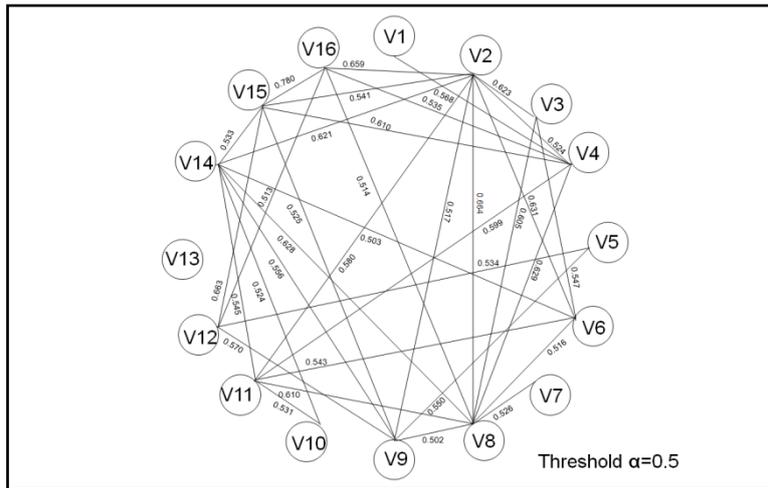


Figure. 13. Correlation Relationship Ring, Threshold $\alpha = 0.5$.

When the threshold was increased gradually, the correlation pattern changed, the correlation lines reduced, as showed in figure 13 and 14.

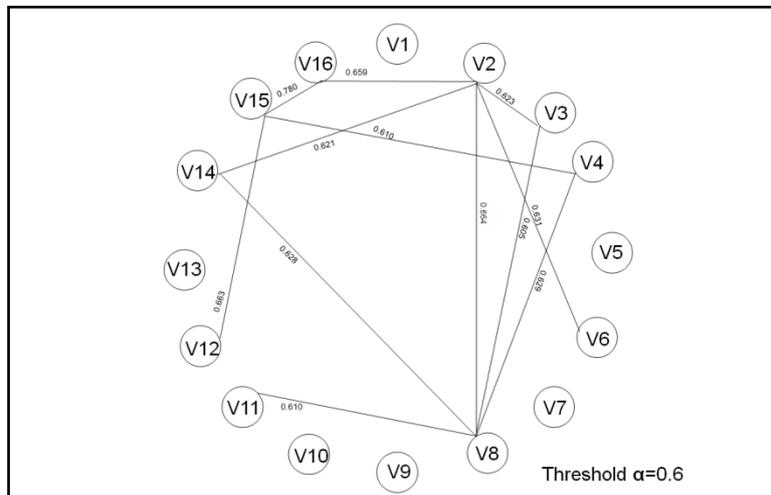


Figure. 14. Correlation Relationship Ring, Threshold $\alpha = 0.6$.

5.6.4 Correlation Relationship Cluster

We cleaned the correlation matrix by setting the threshold value to $\alpha = 0.6$ and $\alpha = 0.66$ in order to create a correlation relationship cluster. In this cluster, we could see that they are group into two strong correlation clusters, as per below figure 15.

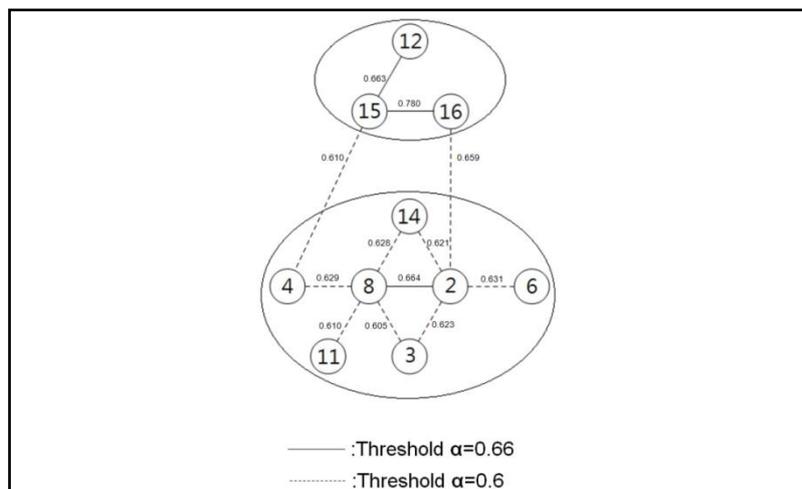


Figure. 15. Correlation Relationship Cluster, Threshold $\alpha = 0.6$, $\alpha = 0.66$.

We cleaned again the correlation matrix by setting the threshold value to $\alpha = 0.66$ in order to create independent correlation relationship clusters.

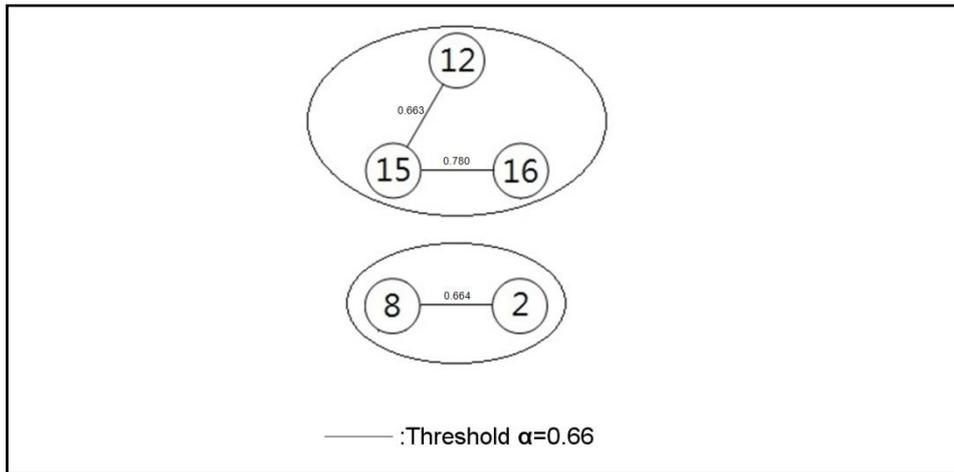


Figure. 16. Correlation Relationship Cluster, Threshold $\alpha = 0.66$.

From figure 16, variable 2 and 8 are integrity and quality assurance. Whereas variable 12, 15 and 16 are strategic support, high bandwidth of communications channels, and microprocessor perform memory and compute capabilities. Because integrity belongs to the category of philosophy, while quality assurance belongs to rule based social system, it supports to the statement that the elements of category philosophy and rule based social system are closely correlated. On the other hand, strategic support belongs to the category of strategic system, while high bandwidth of communications channels, and microprocessor perform memory and compute capabilities belongs to the hardware category, so we could conclude that strategic system and hardware are closely correlated.

5.6.5 Correlation Relationship Tree

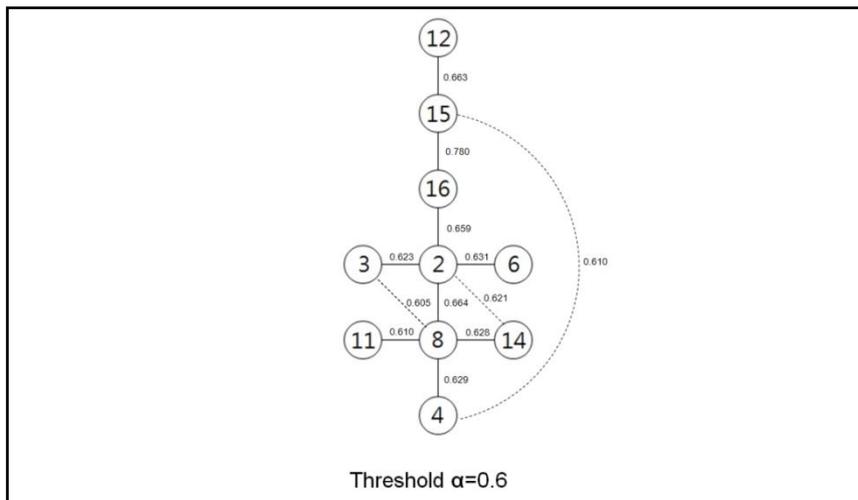


Figure. 17. Correlation Relationship Tree, Threshold $\alpha = 0.6$.

These tree relations enable us to see the main correlation strength, as showed in figure 17. The strongest point was attached between variable 15 (high bandwidth of communication channels) and variable 16 (microprocessor perform memory and compute capabilities). These two variables are from the category of hardware. In conclusion, the items in the hardware category are highly correlated to each other, when a consumer perceived on the hardware supporting the information security system, they perceived them as a package, for they carries the same weight.

VI. CONCLUSION

This paper gives insight into the importance of information security criteria as the main aspect perceived to impact customer trust towards the entire smart grid system. The analysis takes aim at identifying the criteria that could enhance the information security system of a smart grid project and discusses the impact and significance of each of the requirements identified. The major finding of this paper exhibit the agreement from the respondent on the essentiality of the identified sixteen consumer requirements, and the ranking of essentiality of the requirements were generated. The most important consumer requirement of information security in a smart grids system referred to the respondents' respond was privacy concern, followed by confidentiality, integrity, facilities misuse prevention, security in wireless media, availability, quality assurance, reliable systems level, reliable device level, networking issues, strategic support, tactical oversight monitoring system, cryptography and key management, microprocessor perform memory and compute capabilities, mature or proprietary protocols, and high bandwidth of communications channels. Comparison result showed that each of the ranking

was a mixture of from five categories of philosophy, human behavior, rule based social system, strategic system and hardware, and hence each category could be concluded to have a balance value of essentiality.

Histogram test indicated that the average respondents was at least ticking the value of more than average, that was on agreed or strongly agreed side when asked about the importance of identified requirement in conjunction to the information security in a smart grid system.

To prove that all variables are slanted to the right, we made hypothesis and tested it with Chi-Square test. From the result, we concluded that the entire null hypothesis from each class has not been rejected. Also, the results of p_{β} showed that the value is greater than 80% of the entire respondents' agreement on the essentiality of the consumer requirements identified. Therefore, this result supports our null hypothesis that the graphs of all of the variables are slanted to the right.

Individual value plot versus country shows that there are still minority respondents stating that confidentiality is not so important contribute towards the sustainability of an information security in a smart grid system, disagreed where research on a number of related topics was required to further approaches to building advanced protection architecture that can evolve and can tolerate failures, and disagreed that microprocessor performs memory and compute capabilities was an essential element for an information security in a smart grid system. This might due to the rapidly growing economies in different countries have different smart grid infrastructure needs.

Pearson correlation test showed that security in wireless media has very low correlation with others variable. The wireless connections have become an increasingly popular way of providing internet access although it is still in its development process. When the wireless becomes a common method for data transmission especially from the smart meter to the grid, security becomes a highly important part of the wireless network structure. Hence, although the security in wireless media is not correlated to other variable tested, it is still an important asset for all systems using a wireless network.

Employed the output from the Pearson correlation test, we generated two means of ranking for the variables, the lexicographical order to test the essential variables and the correlation relationship ring, cluster and tree to test the essential category. Lexicographical order test showed that integrity, quality assurance and high bandwidth of communications channels has high correlation with each other. Any changes in the unauthorized modification or destruction of information will lead to the changes in the quality assurance performing functions and types of security technologies that should be used across an interface while still meeting that interface's performance requirements. Correlation relationship ring, cluster and tree concluded that the items in the hardware category are highly correlated to each other, when a consumer perceived on the hardware supporting the information security system, they perceived them as a package, for they carries the same weight.

The success of this effort appeared to hinge on utility companies championing information system security initiatives and propagating an awareness of the importance of information security among consumers at all levels of the community. In this study, there was a causal relationship because consumer requirements cause energy authorities to modify their information security policies according to consumer demand.

VII. DISCUSSION

This paper has tutorial contents where some related backgrounds were provided, especially covering the cyber security requirement of smart grid information infrastructure. It provides a combination of methodologies and a set of sixteen identified information security consumer requirements conceptually as original contributions. This paper aims to contribute a sight for the readers to have a conceptual knowledge of the electric power grid and a better understanding of cyber security, and focuses main on helping the government and utilities provider to enhance processes, to develop and prioritize the necessary requirements for securing their information in a smart grid.

ACKNOWLEDGMENT

The authors wish to thank Dr. Siew Hai Yen, Dr. Shu-ichi Kinoshita, Dr. Akiyasu Tomoeda, Tateno Kurina for the comments; and Professor Masayasu Mimura for mentoring. This study was supported by the Meiji University Global COE Program "Formation and Development of Mathematical Sciences Based on Modeling and Analysis", Meiji Institute for Advanced Study of Mathematical Sciences (MIMS), and the Japanese government's Ministry of Education, Culture, Sports, Science, and Technology (MEXT) scholarship.

REFERENCE AND CITATIONS SECTION

- [1] Christoph Weber, Adriaan Perrels, (2000) "Modelling lifestyle effects on energy demand and related emissions", Energy Policy 28, pp. 549-56.
- [2] Inge Røpk, (2001) "New technology in everyday life – social processes and environmental impac", Ecological Economics 38, pp. 403–422.
- [3] R. Stark, H. Hayka, D. Langenberg, (2009) "New potentials for virtual product creation by utilizing grid technology", CIRP Annals - Manufacturing Technology, Volume 58, Issue 1, Pages 143-146.
- [4] Amy Poh Ai Ling, Masao Mukaidono, "Smart Grid Information Security (IS) Functional Requirement", International Journal of Emerging Sciences, Vol.1, No.3, September 2011, pp. 371-386.
- [5] Twiki Research 2010, CNSLab (Computer Networks and Security Lab), available at: <http://cnslab.snu.ac.kr/twiki/bin/view/Main/Research>

- [6] Amy Poh Ai Ling, Masao Mukaidono, "Selection of Model in Developing Information Security Criteria on Smart Grid Security System", Smart Grid Security and Communications, The Ninth International Symposium on Parallel and Distributed Processing with Applications (ISPA), No. 108, May 2011, Korea, pp.91-98; Journal of Convergence, Vol.2, No.1, 2011-6, pp.39-46.
- [7] "Information Security Incident Survey Report", NPO Japan Network Security Association, Security Incident Investigation Working Group, Ver. 1.0, March 31, 2010.
- [8] Jean-Philippe Vasseur, Adam Dunkels, (2010) "Smart Grid", Interconnecting Smart Objects with IP, pp. 305-324.
- [9] Leonardo Meeus, Marcelo Saguan, (2011) "Innovating grid regulation to regulate grid innovation: From the Orkney Isles to Kriegers Flak via Italy", Renewable Energy, Volume 36, Issue 6, June 2011, pp.1761-1765.
- [10] Tony Flick, Justin Morehouse, "Smart Grid: What Is It?", Securing the Smart Grid, 2011, Pages 1-18.
- [11] Ravi Akella, Han Tang, Bruce M. McMillin, "Analysis of information flow security in cyber-physical systems", International Journal of Critical Infrastructure Protection, Volume 3, Issues 3-4, December 2010, Pages 157-173.
- [12] G. Laccetti, G. Schmid, (2007) "A framework model for grid security", Future Generation Computer Systems, Volume 23, Issue 5, June 2007, pp.702-713.
- [13] Simon Perry, "Watt matters – smart grid security", Infosecurity, Volume 6, Issue 5, July-August 2009, Pages 38-40.
- [14] Amy Poh Ai Ling, Mukaidono Masao, "Grid Information Security Functional Requirement Fulfilling Information Security of a Smart Grid System", International Journal of Grid Computing & Applications, Vol. 2, No. 2, June 2011, pp. 1–19.
- [15] Walsham, G., (1996) "The emergence of interpretivism in IS research", Information Systems Research, Vol. 6, pp.376-94.
- [16] Klein, H.K., Myers, M.D., (1999) "A set of principles for conducting and evaluating interpretive field studies in information systems", MIS Quarterly, Vol. 23 pp.67-94.
- [17] Ramberg, Bjørn and Kristin Gjesdal (2005) "Hermeneutics", Stanford Encyclopedia of Philosophy.
- [18] Rossi et al., 1992 R.E. Rossi, D.J. Mulla, A.G. Journel and E.H. Franz, Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecol. Monogr., 62 (1992), pp. 277–314.
- [19] T. Eavis, A. Lopez, RK-hist: an r-tree based histogram for multi-dimensional selectivity estimation, ACM Int. Conf. on Inf. and Knowl. Management, 2007, pp. 475–484.
- [20] Greenwood, P.E., Nikulin, M.S. (1996) A guide to chi-squared testing. Wiley, New York. ISBN 047155779X "Ethics of Information Communication Technology (ICT)", Regional Meeting on Ethics of Science and Technology, Regional Unit for Social & Human Sciences in Asia and the Pacific, UNESCO, access on 16 Sept. 2011, available at http://www2.unescobkk.org/elib/publications/ethic_in_asia_pacific/239_325ETHICS.PDF
- [21] Skiena, S. "Lexicographically Ordered Permutations" and "Lexicographically Ordered Subsets." §1.1.1 and 1.5.4 in Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica. Reading, MA: Addison-Wesley, pp. 3-5 and 43-44, 1990.
- [22] Charl van der Walt, "Introduction to Security Policies, Part One: An Overview of Policies", Introduction to Security Policies, Part One: An Overview of Policies, Nov 2010, available at: <http://www.symantec.com/connect/articles/introduction-security-policies-part-one-overview-policies>
- [23] "Data Leakage Worldwide: Common Risks and Mistakes Employees Make", White Paper, Cisco, 2008 Cisco System, pp. 1-10.
- [24] "Technology Roadmap", Smart Grids, International Energy Agency, Apr. 2011, pp. 1-50.
- [25] "Lobby for smart grid roadmap in Malaysia", Green Purchasing Asia, 2011, available at: <http://www.greenpurchasingasia.com/content/lobby-smart-grid-roadmap-malaysia>
- [26] "Emerging Market Spotlight, South Korea & Taiwan Electronics: Race to the Top", Thomas White, Mac. 2011, In press: <http://www.thomaswhite.com/explore-the-world/emerging-market-spotlight/south-korea-and-taiwan-consumer-electronics.aspx>
- [27] Phongsak Prasithsangaree, "On a framework for energy-efficient security protocols in wireless networks", Computer Communications, Volume 27, Issue 17, 1 November 2004, pp.1716–1729.
- [28] Yuh-Min Tseng, "A resource-constrained group key agreement protocol for imbalanced wireless networks", Computers & Security, Volume 26, Issue 4, June 2007, pp. 331-337.

Modeling and Analysis of Two Wheeler Connecting Rod

K. Sudershn Kumar¹, Dr. K. Tirupathi Reddy², Syed Altaf Hussain³

**(PG student, School of Mechanical Engineering RGM College of Engg. & Technology, Nandyal-518501, India)*

*** (School of Mechanical Engineering, RGM College of Engineering & Technology, Nandyal-518501, India)*

**** (School of Mechanical Engineering, RGM College of Engineering & Technology, Nandyal-518501, India)*

ABSTRACT: The connecting rod is the intermediate member between the piston and the Crankshaft. Its primary function is to transmit the push and pull from the piston pin to the crank pin, thus converting the reciprocating motion of the piston into rotary motion of the crank.

Existing connecting rod is manufactured by using Carbon steel. This paper describes modeling and analysis of connecting rod. In this project connecting rod is replaced by Aluminum reinforced with Boron carbide for Suzuki GS150R motorbike. A 2D drawing is drafted from the calculations. A parametric model of connecting rod is modelled using PRO-E 4.0 software. Analysis is carried out by using ANSYS software.

Finite element analysis of connecting rod is done by considering two materials, viz., Aluminum Reinforced with Boron Carbide and Aluminum 360. The best combination of parameters like Von misses stress and strain, Deformation, Factor of safety and weight reduction for two wheeler piston were done in ANSYS software. Compared to carbon steel, aluminum boron carbide and aluminum 360, Aluminum boron carbide is found to have working factor of safety is nearer to theoretical factor of safety, 33.17% to reduce the weight, to increase the stiffness by 48.55% and to reduce the stress by 10.35% and most stiffer.

Keywords: Connecting rod, Static analysis, Carbon steel, Aluminum, Aluminum reinforced with Boron carbide, Aluminum 360.

I. INTRODUCTION

Connecting rods are widely used variety of engine. The function of connecting rod is to transmit the thrust of the piston to the crank shaft, and as the result the reciprocating motion of the piston is translated into rotational motion of the crank shaft. It consist of a pin –end. A shank section, and crank an end .Pin end and crank end pin holes are machined to permit accurate fitting of bearings. One end of the connecting rod is connected to the piston by the piston pin. Connecting rods are subjected to forces generated by mass and fuel combustion .Theses two forces results in axial load and bending stresses. A connecting rod must be capable of transmitting axial tension, axial compression, and bending stress caused by the thrust and full of the piston and by centrifugal force. Finite element (FEM) Modal is a modern way for fatigue analysis and estimation of the component .The influential component factors are able to change such as material .cross section conditions etc.

In modern automotive internal combustion engine, the connecting rods are most usually made of steel for production engine. But can be made of aluminum or titanium for high performance of engines of cast iron for application such as motor scooters. They are not rigidly

fixed at either end , so that the angle between the connecting rod and piston can change as the rod moves up and down and rotates around the crank shaft .The big end connects to the bearings journal on the throw connecting rod is under tremendous stress from the reciprocating load represented by the piston ,actually stretching and being compressed with every rotation, and the load increases to the third power with increasing engine speed .Connecting rod for automotive applications are typically manufactured by forging from either wrought steel or powder metal. Schematic diagram for connecting rod as shown in figure 1



Fig:1 Schematic diagram of connecting rod.

II. SPECIFICATION OF THE PROBLEM

The objective of the present work is to design and analyses of connecting rod made of Aluminum Reinforced with Boron carbide. Steel and aluminum materials are used to design the connecting rod. In this project the material (carbon steel) of connecting rod replaced with Aluminum Reinforced with Boron carbide. Connecting rod was created in Pro-E. Model is imported in ANSYS 12.0 for analysis. After analysis a comparison is made between exististing steel and aluminum connecting rod viz., Aluminum Reinforced with Boron carbide in terms of weight, factor of safety, stiffnes, deformation and stresse.

III. THEORETICAL CALCULATIONS OF CONNECTING ROD

A connecting rod is a machine member which is subjected to alternating direct compressive and tensile forces. Since the compressive forces are much higher than the tensile force, therefore the cross-section of the connecting rod is designed as a strut and the Rankin formula is used.

A connecting rod subjected to an axial load W may buckle with x-axis as neutral axis in the plane of motion of the connecting rod, {or} y-axis is a neutral axis. The connecting rod is considered like both ends hinged for buckling about x-axis and both ends fixed for buckling about y-axis. A connecting rod should be equally strong in buckling about either axis.

Let

A = cross sectional area of the connecting rod.

L = length of the connecting rod.

σ_c =compressive yield stress.

W_{cr} =crippling or buckling load.

I_{xx} and I_{yy} =moment of inertia of the section about x-axis and y-axis respectively.
 K_{xx} and K_{yy} =radius of gyration of the section about x-axis and y- axis respectively.
 Rankin formula = $(I_{xx}=4I_{yy})$.

3.1 PRESSURE CALCULATION FOR 150CC ENGINE

Suzuki GS 150 R Specifications

Engine type air cooled 4-stroke
 Bore × Stroke (mm) = 57×58.6
 Displacement = 149.5CC
 Maximum Power = 13.8bhp@8500rpm
 Maximum Torque = 13.4Nm@6000rpm
 Compression Ratio = 9.35/1
 Density of Petrol C_8H_{18} = 737.22kg/m^3
 = $737.22\text{E}^{-9}\text{kg/mm}^3$
 Temperature = 60F = 288.855K
 Mass = Density × Volume
 = $737.22\text{E}^{-9} \times 149.5\text{E}^3$
 = 0.11Kg
 Molecular Weight of Petrol 114.228 g/mole
 From Gas Equation,
 $PV=Mrt$
 $R = R^*/Mw = 8.3143/114.228$
 = 72.76
 $P = (0.11 \times 72.786 \times 288.85) / 149.5\text{E}^3$
 $P = 15.469 \text{ Mpa}$.

3.2 DESIGN CALCULATION FOR CARBON STEEL

Thickness of flange & web of the section = t
 Width of section $B = 4t$
 The standard dimension of I SECTION.

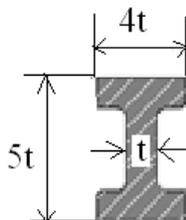


Figure2: Standard dimension of I - Section.

Height of section $H=5t$

Area of section $A = 2(4t \times t) + 3t \times t$

$A = 11t^2$

MI of section about x axis:

$$I_{xx} = 1/12 (4t (5t)^3 - 3t (3t)^3) = 419/12 t^4$$

MI of section about y axis:

$$I_{yy} = (2 \times 1/12 t (4t)^3 + 1/12 (3t) t^3) = 131/12 t^4$$

$I_{xx} \setminus I_{yy} = 3.2$

Length of connecting rod (L) = 2 times the stroke

$$L = 117.2 \text{ mm}$$

Buckling load $w_B = \text{maximum gas force} \times \text{F.O.S}$

$$W_B = \frac{\sigma_c \times A}{1 + a(L \setminus K_{xx})^2} = 37663 \text{ N}$$

$\sigma_c = \text{compressive yield stress} = 415\text{MPa}$

$K_{xx} = I_{xx} \setminus A$

$K_{xx} = 1.78t$

$a = \sigma_c \setminus \pi^2 E$

$a = 0.0002$

By substituting $\sigma_c, A, a, L, K_{xx}$ on W_B then

$$4565t^4 - 37663t^2 - 81639.46 = 0$$

$$t^2 = 10.03$$

$$t = 3.167\text{mm}$$

$$t = 3.2\text{mm}$$

Width of section $B = 4t = 12.8\text{mm}$

Height of section $H = 5t = 16\text{mm}$

$$\text{Area } A = 11t^2 = 112.64\text{mm}^2$$

Height at the big end (crank end) = $H_2 = 1.1H$ to $1.25H$

$$H_2 = 17.6\text{mm}$$

Height at the small end (piston end) = $0.9H - 0.75H$

$$H_1 = 12\text{mm}$$

3.3 DESIGN CALCULATION FOR ALUMINUM 360.

Buckling load $w_B = \text{maximum gas force} \times \text{F.O.S}$

$$W_B = \frac{\sigma_c \times A}{1 + a(L \setminus K_{xx})^2} = 31386 \text{ N}$$

$\sigma_c = \text{compressive yield stress} = 172 \text{ MPa}$

$K_{xx} = I_{xx} \setminus A$

$K_{xx} = 1.78t$

$a = \sigma_c \setminus \pi^2 E$

$a = 0.002$

By substituting $\sigma_c, A, a, L, K_{xx}$ on W_B then

$$4565t^4 - 37663t^2 - 81639.46 = 0$$

$$t^2 = 16.64$$

$$t = 4.08\text{mm}$$

$$t = 4.1\text{mm}$$

Width of section $B = 4t = 16.4\text{mm}$

Height of section $H = 5t = 20.5\text{mm}$

$$\text{Area } A = 11t^2 = 184.91\text{mm}^2$$

Height at the big end (crank end) = $H_2 = 1.1H$ to $1.25H$

$$H_2 = 25.625\text{mm}$$

Height at the small end (piston end) = $0.9H - 0.75H$

$$H_1 = 18.45\text{mm}$$

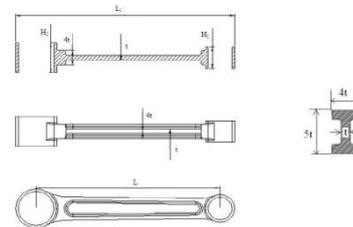


Figure3: 2D drawing for connecting rod.

IV. SPECIFICATION OF EXISTING CONNECTING ROD

Table 1 shows the specifications of the connecting rod for carbon steel (Suzuki GS). The typical chemical composition of the material is 0.61% C, 0.095% Al, 0.82% Mn, 0.00097% Br, 0.145% C, 7.8Co, 75.56Fe and 3.25 Mo.

Table: 1 Specifications of connecting rod:

S.No	Parameters	Value
1	Bore × Stroke (mm)	57×58.6
2	Length of connecting rod	112 mm
3	Thickness of connecting rod	For C.S = 3.2mm For AL 360 = 4.1 mm
4	Width of connecting rod	For C.S = 12.8mm For AL 360 = 16.4 mm

5	Height of connecting rod	For C.S H ₁ =12mm H ₂ =17.6mm
		For AL 360 H ₁ =18.45mm H ₂ =25.625mm

V. STRUCTURAL ANALYSIS OF CONNECTING ROD

Dimensions of Width and height of the connecting rod is For C.S = 12.8mm and For AL 360 = 16.4 mm. A 3-D model of connecting is used for analysis in ANSYS 12.0. The loading conditions are assumed to be static. Analysis done with pressure load applied at the piston end and restrained at the crank end or other load applied at the crank end and restrained at the piston end. The element chosen is SOLID 187, it was used with the tetrahedral option, making it a 10-node element with 3 degrees of freedom at each node. The finite element analysis is carried out on carbon steel connecting rod as well as on three different materials of carbon steel, aluminum boron carbide and aluminum 360. From the analysis the equivalent stress (Von-mises stress), strain, displacements were determined and are shown in figure 4-15. Table 2 shows the comparative of factor of safety for three different materials.

VI. STRUCTURAL ANALYSIS OF CONNECTING ROD

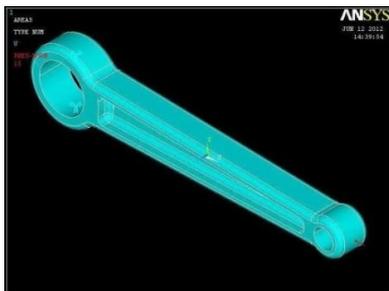


Figure4: loads and boundary conditions.

6.1 CARBON STEEL:

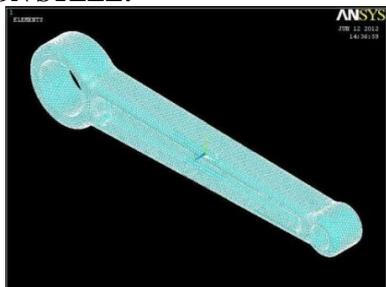


Figure5: Mesh of carbon steel.

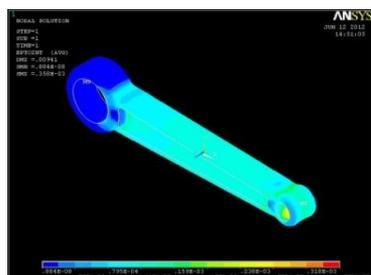


Figure6: von mises strain of carbon steel.

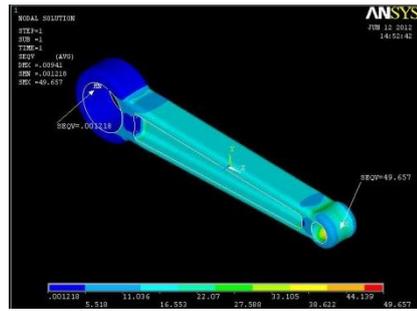


Figure7: von mises Stress for carbon steel.

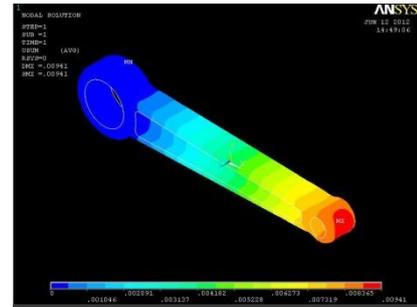


Figure8: Displacement of carbon steel.

6.2 ALUMINUM 360:

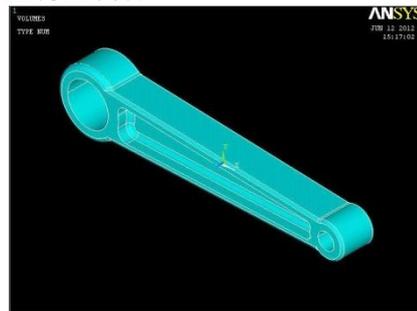


Figure9: modeling of aluminum 360.

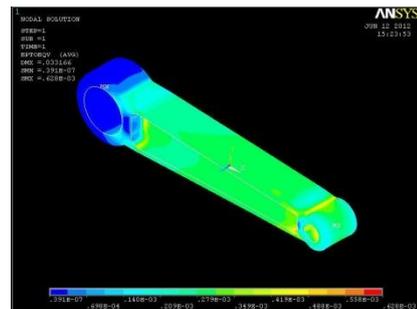


Figure10: von mises strain of aluminum 360.

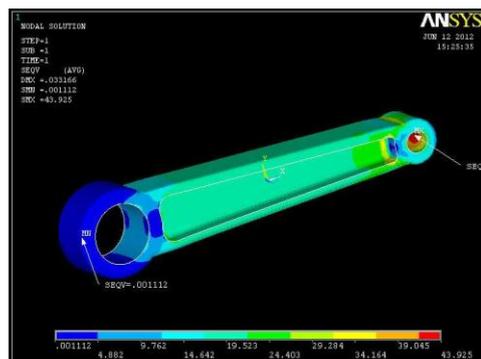


Figure11: von mises stress of aluminum 360.

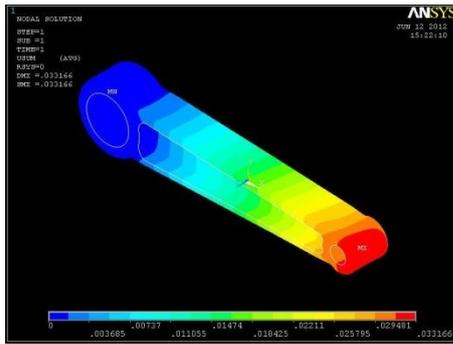


Figure12: Displacement of aluminum 360.

6.3 ALUMINUM BORON CARBIDE:

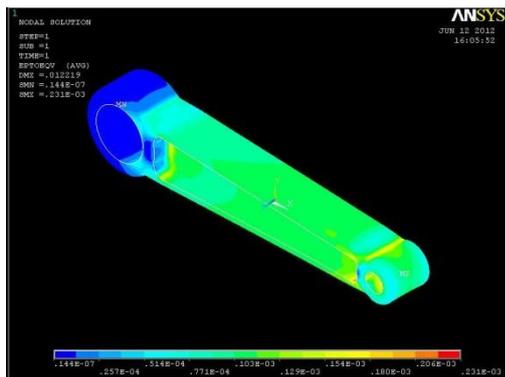


Figure13: von mises train of aluminum boron carbide.

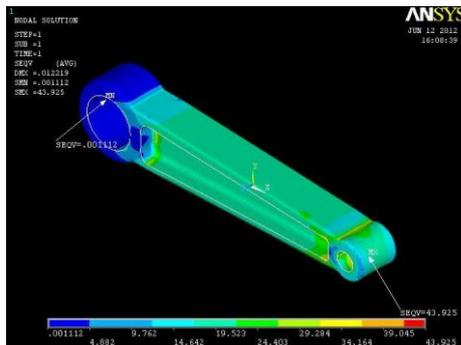


Figure14: von mises stress aluminum boron carbide.

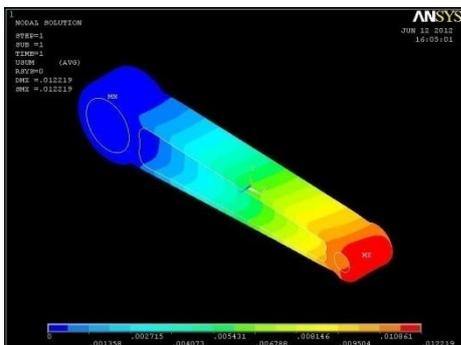


Figure15: Displacement aluminum boron carbide.

VII. RESULTS

Table 2. Comparison of factor of safety.

Material properties	Carbon steel	Aluminum 360	Aluminum boron carbide
Yield strength	415	172	300

(Mpa)			
Tensile strength (Mpa)	540	317	485
Theoretical Factor of safety (N)	6	6	6
Allow stress (Mpa)	69.16	28.47	50
Anslys result (Mpa)	49	43.925	43.925
Working Factor of safety (N)	8.47	4	6.95

7.1. RESULT FOR WEIGHT OF CONNECTING ROD

Density of Carbon steel = $7.87 \times 10^{-6} \text{ Kg/mm}^3$
 Volume of connecting rod = 92419.78 mm^3
 Weight of connecting rod = Density \times Volume
 = $7.87 \times 10^{-6} \times 92419.78$
 = 0.727 kg

1. Density of Al 360 = $2.685 \times 10^{-6} \text{ kg/mm}^3$
 Volume of connecting rod = 180935.21 mm^3
 Weight of connecting rod = Density \times Volume
 = 0.48581 kg

Percentage of reduction in weight = $\frac{W \text{ of Carbon steel} - W \text{ of Al 360}}{W \text{ of Carbon steel}}$
 = $\frac{0.727 - 0.48581}{0.727}$
 = 0.3317

Aluminum boron carbide = $\frac{W \text{ of Carbon steel} - W \text{ of Aluminum boron carbide}}{W \text{ of Carbon steel}}$
 = $\frac{0.727 - 0.48581}{0.727}$
 = 0.3317

7.2. RESULT FOR STIFFNESS OF CONNECTING ROD:

Carbon steel
 Weight of connecting rod = 0.727Kg
 Deformation = 0.00941mm
 Stiffness = Weight/Deformation
 = $\frac{0.727}{0.0094}$
 = 77.34 kg/mm

Aluminum360
 Weight of connecting rod = 0.48581Kg
 Deformation = 0.0033166mm
 Stiffness = Weight/Deformation
 = $\frac{0.48581}{0.00331}$
 = 146.77 kg/mm

Aluminum boron carbide
 Weight of connecting rod = 0.48581Kg
 Deformation = 0.012219mm
 Stiffness = Weight/Deformation
 = $\frac{0.48581}{0.012219}$
 = 39.7585 kg/m

7.3. RESULT FOR PERCENTAGE OF INCREASE IN STIFFNESS:

Aluminum 360 = $77.34 - 146.77 / 77.34$
 = -0.8977
 Aluminum boron carbide = $77.34 - 39.7585 / 77.34$
 = 0.4855

7.4. RESULT FOR PERCENTAGE OF STRESS REDUCTION:

Aluminum 360 = $49.625 - 43.925 / 49.625$
 = 0.1035
 Aluminum boron carbide = $49 - 43.925 / 49$
 = 0.1035

VIII. GRAPHS

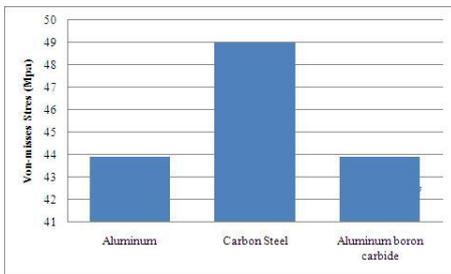


Fig 16: Von-Misses Stress for three materials.

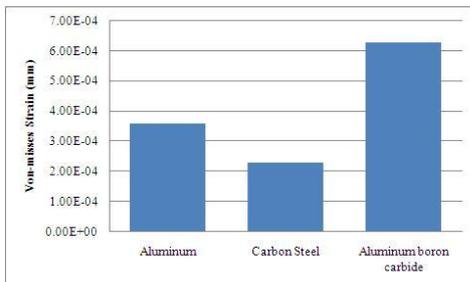


Fig 17: Von-Misses Strain for three materials.

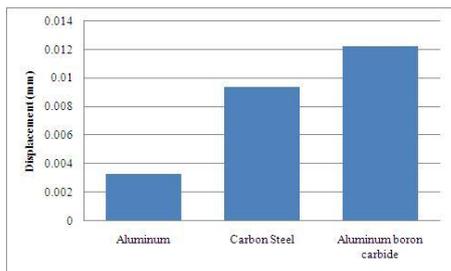


Fig 18: Displacement for three materials.

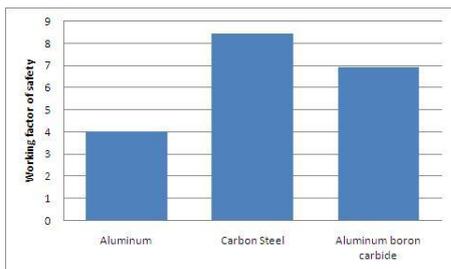


Fig 19: working factor of safety for three materials.

IX. CONCLUSIONS

By checking and comparing the results of materials in above tables and finalizing the results are shown in below.

For considering the parameters,

The working factor of safety is nearer to theoretical factor of safety in aluminum boron carbide.

Percentage of reduction in weight is same in Aluminum 360 and aluminum boron carbide.

Percentage of increase in stiffness in aluminum boron carbide is more.

Percentage of reducing in stress ALUMINIUM BORON CARBIDE and ALUMNUM is same than CARBON STEEL.

REFERENCES

1. Afzal, A. and A. Fatemi, 2004. "A comparative study of fatigue behaviour and life predictions of forged steel and PM connecting rods". SAE Technical Paper, 1: 1529.
2. Anonymous, 2008. "Nissan Z24engine Maintenance and Repayments catalogue". Megamotor Co.
3. Chen, N., L. Han, W. Zhang and X. Hao, 2006. "Enhancing Mechanical Properties and Avoiding Cracks by Simulation of Quenching Connecting Rod". Material Letters, 61: 3021-3024.
4. El-Sayed, M.E.M. and E.H. Lund, 1990. "Structural optimization with fatigue life constraints," Engineering Fracture Mechanics, 37(6): 1149-1156.
5. Jahed Motlagh, H., M. Nouban and M.H. Ashraghi, 2003. "Finite Element ANSYS". University of Tehran Publication, PP: 990.
6. Khanali, M., 2006. "Stress analysis of frontal axle of JD 955 combines". M.Sc. thesis. Thran University, 124.
7. Kolchin, A. and V. Demidov, 1984. "Design of Automotive Engines". MIR Publication.
8. Meriam, J.L. and L.G. Kraige., 1998. Engineering Mechanics, 5th Edition, New York, john wiley, 712.
9. Reppen, B., 1998. "Optimized Connecting Rods to Enable Higher Engine Performance and Cost Reduction," SAE Technical Paper Series, Paper No. 980882.
10. Shigley, J.E. and C.R. Mischke, 2001. "Mechanical Engineering Design", McGraw-Hill, New York, 776.
11. Webster, W.D., R. Coffell and D. Alfaro, 1983. "A Three Dimensional Finite Element Analysis of a High Speed Diesel Engine Connecting Rod," SAE Technical Paper Series, Paper No. 831322.

BOOKS:

1. Machine design by R.S. KHURMI, J.K GUPTA.
2. Design data by PSG.
3. A text book of Machine Design by S.Md. JALALUDEEN.

Effect of CETANE Improver Additives on Emissions

K. Velmurugan¹, S. Gowtham²

¹M.E-Internal combustion engineering, ²M.E-Internal combustion engineering
Velammal engineering college, Chennai-600053.

ABSTRACT: Exhaust gases of an engine can have upto 2000 ppm of oxides of nitrogen. Most of this will be nitrogen oxide (NO), with a small amount of nitrogen dioxide (NO₂). NO_x is very undesirable. Regulations to reduce NO_x emissions continue to become more and more stringent year by year. Released NO_x reacts in the atmosphere to form ozone and is one of the major causes of photochemical smog. NO_x is created mostly from nitrogen in the air. Nitrogen can also be found in fuel blends. At high temperature and pressure higher levels of NO_x is created and at low temperature lower level of NO_x is produced. In addition to temperature, the formation of NO_x depends on pressure and air-fuel ratio. Engine emission from diesel fuel, paying special attention to the most concerning emissions. Oxides of nitrogen is not only in mass and composition but also in size distributions. Many of the design changes for reduction in NO_x emissions result in higher brake specific fuel consumption. Test is carried out in a single cylinder direct injection diesel engine. Cetane number describes the auto ignition quality of the diesel fuel. Cetane improver additive of neopentane is used with the varying proportions of 1ml, 3ml, 5ml to the diesel fuel respectively. Addition of cetane improver additive to the diesel fuel is cost effective way to control NO_x emission. Diesel fuel with the 3ml additive of neopentane shows the significant reduction in NO_x and smoke. The sensitivity of NO_x to change in cetane number is higher at low load than at high load. It is found that NO_x emissions were reduced at low load than at high load.

KEYWORDS: NO_x, HC, CO, Cetane additive.

I. INTRODUCTION

Emissions from automobiles were identified for the first time as a major contribution to urban air pollution during 1950's. Photochemical reactions between unburned hydrocarbons and nitrogen oxides emitted mostly by vehicles were found responsible for production of a host of secondary pollutants including ozone and other oxidants in atmosphere. The 'brown haze' appearing above the ground level in the Los-Angeles area as a result of the photochemical reactions between the vehicular air pollution, the amount of emissions depending largely on their design, operating conditions and the characteristics of the fuel.

The increased industrialization and motorization of the world in recent years has resulted in great demand for petroleum products. Petroleum is the largest single source of energy which has been consuming by the world's population, exceeding the other energy resources such as natural gas, coal, nuclear and renewable. 90% of energy consumption of the world is from petroleum fuels. Petroleum based natural gas, coal, nuclear and renewable. Petroleum based fuels are obtained from limited reserves and estimated to last only for new decades.

According to international Energy Outlook 2007 published by the Energy Information Administration, the world consumption for petroleum and other liquid fuel will grow from 83 million barrels/day in 2004 to 97 million barrels/day in 2015 and just over 118 million barrels/day in 2025 under these growth assumptions, approximately half of the world's total resources would be exhausted by 2025. Therefore the future energy availability is a serious problem for us.

Major global concern is environmental concern or climate change such as global warming. Global warming is related with the green house gases which are mostly emitted from the combustion of petroleum fuels. The Intergovernmental Panel on Climate Change (IPCC) concludes in the Climate change 2007 that, because of

global warming effect the global surface temperatures are likely to increase 1.1C to 6.4C between 1990 and 2100.

II. CETANE NUMBER & ADDITIVE

Cetane number or CN is a measurement of the combustion quality of diesel fuel during compression ignition. It is a significant expression of diesel fuel quality among a number of other measurements that determine overall diesel fuel quality.

Cetane number or CN is actually a measure of a fuel's ignition delay; the time period between the start of injection and the first identifiable pressure increase during combustion of the fuel. In a particular diesel engine, higher cetane fuels will have shorter ignition delay periods than lower cetane fuels. Cetane numbers are only used for the relatively light distillate diesel oils.

III. TYPICAL VALUES

Generally, diesel engines run well with a CN from 40 to 55. Fuels with higher cetane number which have shorter ignition delays provide more time for the fuel combustion process to be completed. Hence, higher speed diesels operate more effectively with higher cetane number fuels. There is no performance or emission advantage when the CN is raised past approximately 55; after this point, the fuel's performance hits a plateau. Dimethyl ether may prove advantageous as a future diesel fuel as it has a high cetane rating (55) and can be produced as a biofuel. Alkyl nitrates (2-ethyl hexyl nitrate) and di-tert-butyl peroxide are used as additives to raise the cetane number.

CETANE ADDITIVE:

Generally, cetane additive is used in the diesel engine for controlling NO_x emissions. There are certain cetane additives used widely, as Ethyl hexyl nitrate, alkyl nitrate, peroxide compounds, methyl oleate. In this paper Di-MethylPropane (DMP) or Neopentane is used as the

cetane additive. DMP is a highly volatile liquid on a cold day in an ice bath, or when compressed to high pressure. DMP is an extremely flammable gas at room temperature. When DMP is added to the canola methyl ester it will result in better auto-ignition of the fuel and low NO_x emission. DMP is added as 1ml,3ml,5ml to the diesel fuel.

PROPERTIES OF DMP:

Boiling point :9.5°C
 Molar mass :72.15 g mol⁻¹
 Melting point : -17°C,256 K,1°F
 Exact mass : 72.0939900384 g mol⁻¹
 Chemical formula:C₅H₁₂

DIESEL ENGINE POLLUTANTS

The pollutants from diesel fuel vehicles are Particulate Matter (PM), smoke, NO_x, Sulphur di-oxide, CO and HC. Most pollutants are emitted from the exhaust. Because diesel engines operate at high air-fuel ratios, they tend to have low HC and CO emissions. They have considerably higher PM emissions than gasoline-fueled vehicles; however, for heavy-duty vehicles CO, HC and NO_x emissions in the exhaust also vary with driving modes, engine speed and load.

POLLUTANT FROM DIESEL ENGINE

Smoke, Oxides of Nitrogen,Particulate matter, Oxides of Sulphur and Hydro carbon.

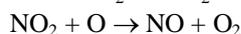
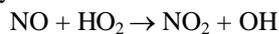
FOLLOWING FACTORS CAUSE THE POLLUTANTS

- ◆ Incomplete combustion,
- ◆ Injection of fuel
- ◆ Air-fuel ratio
- ◆ Time of injection
- ◆ High excess air
- ◆ Availability of oxygen
- ◆ Fuel atomization
- ◆

IV. MECHANISM OF NO_x FORMATION

One of the primary diesel engine pollutants, NO_x is formed by reaction between oxygen and nitrogen in the combustion chamber. NO_x formation dramatically increases with increasing combustion temperature, combustion efficiency, and combustion pressure.

NO_x emissions can be reduced by reducing power, lowering the intake air temperature, retarding the injector timing, reducing the coolant temperature, and/or reducing the combustion temperature. These often reduce fuel economy.



During the "premixed" or uncontrolled diesel combustion phase immediately following the ignition delay, fuel-air mixture with spread in composition about stoichiometric burns due to spontaneous ignition and flame propagation.

CARBON MONOXIDE

Diesels, however, always operate well on the lean side of stoichiometric air fuel ratio CO emissions from diesels

are low enough to be unimportant, therefore, and will not be discussed further.

HYDROCARBON FORMATION MECHANISM

Diesel fuel contains hydrocarbon compounds with higher boiling points, and hence higher molecular weights, than gasoline. Also, substantial pyrolysis of fuel compounds occurs within the fuel sprays during the diesel combustion process.

Hydrocarbon emission levels from diesels vary widely with operating conditions, and different HC formation mechanisms are likely to be most important at different operating modes. Engine idling and light-load operation produce significantly higher hydrocarbon emissions than full-load operation.

Wall temperatures affect HC emissions, suggesting that wall quenching is important, and under especially adverse conditions very high cyclic variability in the combustion process can cause an increase in HC due to partial burning and misfiring cycles.

PARTICULATE FORMATION MECHANISM

Diesel particulates consist principally of combustion generated carbonaceous material (soot) on which some organic compounds has become adsorbed. Most particulate material results from incomplete combustion of fuel hydrocarbon; the lubricating oil contributes some. The composition of particulate matter depends on the conditions in the engine exhaust and particulate collection system. At temperatures above 500°C, the individual particles are principally clusters of many small spheres or spherules of carbon (with a small amount of hydrogen) with individual spherule diameters of about 15 to 30 nm. As temperatures decrease below 500°C, the particles become coated with adsorbed and condensed high molecular weight organic compounds which include: unburned hydrocarbons, oxygenated hydrocarbons. The condensed material also includes inorganic species such as sulfur dioxide, nitrogen dioxide and sulfuric acid.

V. EXPERIMENTAL PROCEDURE

The engine was allowed to run with sole fuel at various loads for nearly 10 minutes to attain the steady state and constant speed conditions. Then the following observations were made.

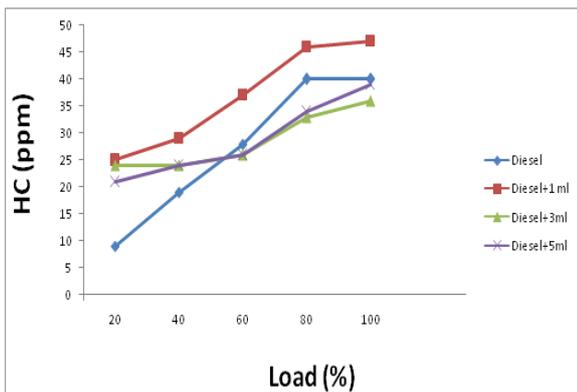
1. The water flow was maintained constant throughout the experiment.
2. The load, speed and temperature indicators were switched on.
3. The engine was started by cranking after ensuring that there is no load.
4. The engine was allowed to run at the rated speed of 1500 rev/min for a period of 20 minutes to reach the steady state.
5. The fuel consumption was measured by a stop watch.
6. Smoke readings were measured using the Hartridge Smoke meter at the exhaust outlet.
7. The amount of NO_x was measured using exhaust gas analyzer.
8. The exhaust temperature was measured by using a sensor.

9. Then the load was applied by adjusting the knob, which was connected to the Eddy Current Dynamometer.
10. Load is applied gradually by 20,40,60,80,100 respectively.
11. Each readings of the load is noted .
12. First, diesel in used in the engine and the reading are noted.
13. The engine is run by diesel with adding 1ml of DMP.
14. The engine is run by diesel with adding 3ml of DMP.
15. The engine is run by diesel with adding 5ml of DMP.
16. Experiment is conducted using sole fuel and sole fuel with additive of Neopentane.
17. From the readings of Added DMP, without DMP graph is plotted based on the tabulated reading.

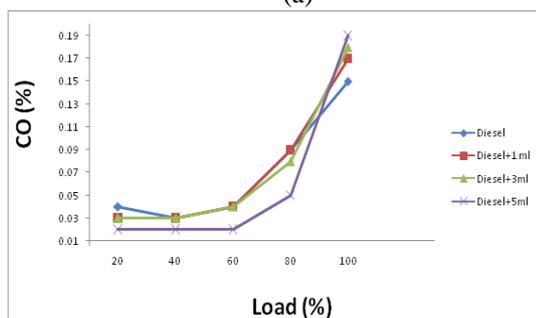
VI. RESULTS AND DISCUSSION

Fig (a) shows that the variation in hydrocarbon emission with time for different loads of sole diesel fuel and diesel fuel with additive of Neopentane . From low load to high load application HC emission is increased .At high load condition the friction power found to be higher than no load condition. At high load condition more amount of fuel is admitted.HC emission is remarkable in the idling and low load because of high heat of vaporization of the fuel.

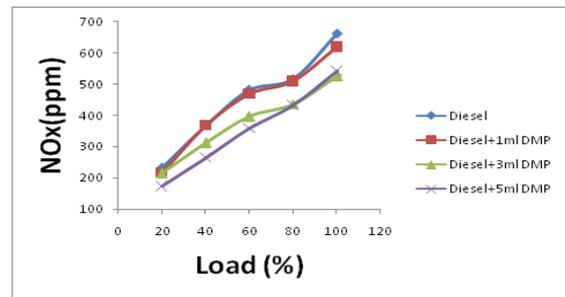
Fig(b) shows that the variation in Carbon monoxide emission with time for different loads of diesel fuel and diesel fuel with additive of Neopentane.CO concentration is lower at the idling and low load and it is slightly increased at high load because an increased premixed combustion due to o longer ignition delay.



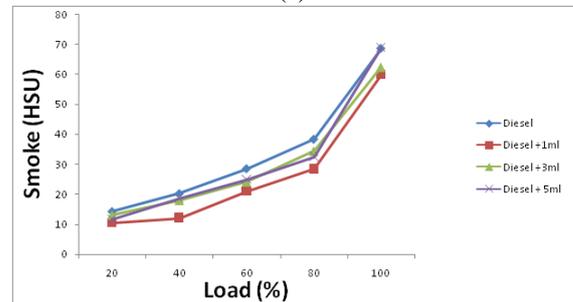
(a)



(b)



(c)



(d)

Fig(c) shows that the variation in Oxides of nitrogen emission with time for different loads of diesel fuel and diesel fuel with additive of Neopentane. NOx emission increases with the increase of load and slightly increases with the addition of cetane improver. Due to the high engine operating temperature at high load NOx is slightly increased. In addition to temperature formation of NOx depends on pressure and air-fuel ratio.

Fig(d) shows that the variation in Smoke emission with time for different loads of diesel fuel and diesel fuel with additive of Neopentane. Carbon spheres are generated in the combustion chamber in the fuel rich zones where there is not enough oxygen to convert all carbon to CO₂. Because of the high compression ratios of CI engines, a large expansion occurs during the power stroke. A single soot particle may contain upto 5000 carbon spheres.

VII. CONCLUSION

Emission characteristics of the single cylinder diesel engine with DMP or Neopentane as an Cetane improver additive were investigated with different load conditions. The result of the study are summarized as follows:

1. From the emission analysis it is observed that there is a 21% reduction in NOx emission with the 3ml of DMP at full load, without increasing the smoke level.
2. HC, CO emission is reduced by 17%, 20% respectively.
3. From the above analysis, it is found that 3ml of DMP with diesel fuel produces significant reduction in NOx, HC, CO emission than the sole diesel fuel, 1 ml of DMP with diesel fuel, 5 ml of DMP with diesel fuel.
4. 0.003% of DMP is selected as the suitable proportion of a cetane improver for the significant reduction of emissions in the single cylinder diesel engine.

REFERENCES

1. K.Varatharajan, M.Cheralathan, R.Velraj.(2011). Mitigation of NOx emissions from a jatropha biodiesel fuelled DI diesel engines using anti oxidant additives.
2. Gvidonas Labeckar, Stasys Slavinskas (2005). The effect of rapeseed oil methyl ester on direct injection diesel engine performance and emissions.
3. Jiafeg Sun, Jerald A.Caton, Timothy J.Jacobs(2010). Oxides of nitrogen emissions from biodiesel fuelled diesel engines.
4. E.Rajasekar, A.Murugesan, R.Subramanian, N.Nedunchezian(2010). Review of NOx reduction technologies in CI engines fuelled with oxygenated biomass fuels.
5. Starr, M.E.(2003). Influence on transient emissions at various injection timings, using cetane improvers, biodiesel, and low aromatic fuels.
6. Nandi, M. Jacobs, D. C. Kesling, H. S. Liotta, F. J.(1994). The Performance of a Peroxide based Cetane improvement Additive in Different Diesels Fuels.
7. Stocky. T.(1994). Role of Cetane Improvers in reformulated Diesel .
8. Liotta. F. J.(1993). A Peroxide Based Cetane Improvement Additive with Favorable Fuel Blending Properties.
9. Schwab, S.D., G.H. Guinther, T.J. Henly, K.T. Miller.(1999). The effects of 2-ethylhexyl nitrate and di-tertiary-butyl peroxide on the exhaust emissions from a heavy-duty diesel engine.
10. Toshitaka Minami, Kouichi Takeuchi, Naoki Shimazaki(1995). Reduction of diesel engine NOx emission using cetane improvers with pilot injection.
11. Ismet celikten, Atilla koca, Mehmet Ali Asslan(2009). comparison of performance and emission of diesel fuel ,rapeseed and soyabean oil methyl esters injected at different pressures.
12. Magentas lapuerta, Octavia Arma, Jose Rodriguez-Fernandez(2008). Effect of biodiesel fuels on diesel engine emissions.
13. Marina Kousoulidou, georgios Fontaras(2010). Biodiesel blend effects on common-rail diesel combustion and emissions.
14. Charles L.Peterson, Craig Chase(1994). ethyl ester of rapeseed used as a biodiesel fuel.
15. Gvidonas Labeckar, Stasys Slavinskas(2007). Study of exhaust emissions of direct injection diesel engine operating on ethanol, petrol and rapeseed oil blends.
16. Cherg-Yuanlin, Kuo-huawary(2003). Diesel engine performance and emission characteristics using three phase emulsions as fuel.
17. Laureano canoira, Ramon Alcanta, Susana Torcal(2006). Nitrated biodiesel as a cetane number enhancer.
18. Mustafa Ertanc Tat(2011). Cetane number effect on the energetic and exergetic efficiency of a diesel engine fuelled with biodiesel.
19. Constantine Arcoumanis, choongsic Base, Roy Crookes (2007). the potential of di-methyl ether (DME) as an alternative fuel for Compression Ignition engines.

Biometrics Security System using the Smart System concept

Nivetha. R¹, Shanmuga Priya. S²

^{1,2}(Department of information technology, Sri Manakula Vinayagar Engineering College, Puducherry

ABSTRACT: The Biometrics field has got much advancement each and every year. Each advancement aims at promoting a high level of security when compared to the previous model in that field. Our paper deals with providing a high level of security as well as efficient authentication system by using the "smart system" concept. According to the smart system concept, we have coupled the two existing efficient systems of biometrics such as the "Iris scanning" and the "Finger print recognition" systems.

Keywords: "CCD Camera, Finger Print Recognition System, Iris Scan and Smart System".

I. INTRODUCTION

In every biometrics systems there is some kind of drawbacks [1][7][8][9]. Our paper aims to overcome this in an efficient manner by using the smart system concept. Though biometric security is ahead of all other security systems, still intruders and hackers deprive its functionality. Improvising security just with the single body part seems to be lacking effectiveness. Hence, in this paper we have tried to devise a new security system in which both IRIS SCAN and FINGER PRINT recognition are combined to prove its worth. It's an attempt to couple the existing technologies rather than to go in for search of a new one. Based on our idea the implementation will replace the traditional ID methods such as P.I.N numbers for accessing ATMs and virtually every other electronic device used for conducting business where identification is a requirement and pre requisite. The smart system deals with combining IRIS SCAN and the FINGER PRINT Recognition Technologies and use the combined value as the security input data to provide security.

This can be done by following the two phases:

1. Registration Phase
2. Authentication Phase

II. REGISTRATION PHASE

The Registration phase is done to the new individual who wants to gain access to the system. The Registration phase is the first phase where the individual undergoes an Iris scan and then the Finger print recognitions. The person's Iris is scanned images and the Finger print images are taken, stored in the security database for the first time when they comes for registration.

2.1 THE IRIS SCAN

IRIS: The **iris** (plural: *irides* or *irises*) is a thin, circular structure in the eye, responsible for controlling the diameter and size of the pupils and thus the amount of light reaching the retina. "Eye color" is the color of the iris, which in humans can be green, blue, or brown. [2] The image of the scanned Iris of the human eye is shown in the figure1.

The Iris Scanning is done with help of a CCD camera. CCD is abbreviated as charge-coupled-device. The person is

made to stand 13 to 15'' from a CCD camera. The CCD camera captures very-high resolution of the images and the accuracy is also very high. Human Iris has 260 independent variables so no two individuals can have the same kind of Iris structure.

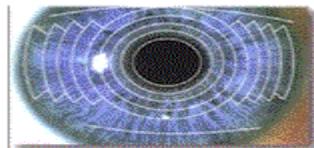


Figure1. Iris scanned image

2.2 FINGER PRINT RECOGNITION

MINUTIAE: The small lines on the surface of the finger. The individual places his/her finger over the glass plate, which resides over the high-resolution CCD camera. These images are relatively large in size and so there can be many permutations and combinations to adjust the threshold value. [2][3][4][5][6] This technology is neither too expensive, nor does it require extensive user – training. It's also simple to implement since a fingerprint reader can sit on a mouse or a keyboard or simply connect like one. It also gives a high level accuracy with 35 independent variables. The finger print image is shown in the figure2. The image captured is compared to that in the system's database. Positive identification or rejection, as the case may be, is based on the number of minutiae (small lines on the surface of the skin) that match. The number that must match for the result to be 'hit' or 'no hit' can be defined. After the Finger print is scanned it is stored in the security database as it was done in the Iris scan. Then the Smart system concept is done by coupling the Iris and the Finger print scanned images.



Figure2. Finger print with minutiae

III. SMART SYSTEM CONCEPT

The concept of the security system is to combine the Iris and the Finger print value from the smart card. It follows the below mentioned algorithm.

Step1: Take the Iris variable as Iris_val.

During the registration of the concerned person into the security, his iris image has to be captured by a high – end CCD camera and then the captured image is stored in the database and in consequent with that an iris variable is generated. (number ranging from 0 to 9).

Step2: Take the Finger print variable as Finger_val.

The recorded image has to be converted as a unique number as follows, Positive identification or rejection, as the case may be, is based on the number of minutiae (small lines on the surface of the skin) that match.

Then a number is generated as follows

- Let X is the number of minutiae in contact with the surface and Y is the number of left out surface.
- Then the number Finger_val is given as XY (number varying from 0 to 9).

Step3: Couple the Iris_val and the Finger_val.

Consider for Example

The Iris_val is generated as 13579 and The Finger_val is generated as 24681. Then these two values are coupled and this coupled value (Couple_val) is used in the smart system to provide security. The two values are coupled as shown below.

$$\text{Iris_val} = 13579 \text{ (x1x2x3x4x5)}$$

$$\text{Finger_val} = 24681 \text{ (y1y2y3y4y5)}$$

$$\text{Couple_val} = 1234567891 \text{ (x1y1x2y2x3y3x4y4x5y5)}$$

Where x1x2x3x4x5 are the digits of the Iris_val and y1y2y3y4y5 are the digits of the Finger_val. These two values are cross coupled and the coupled value obtained consist of x1y1x2y2x3y3x4y4x5y5 which are the digits of the Iris_val and the Finger_val. The coupling process is represented diagrammatically in Figure 3.

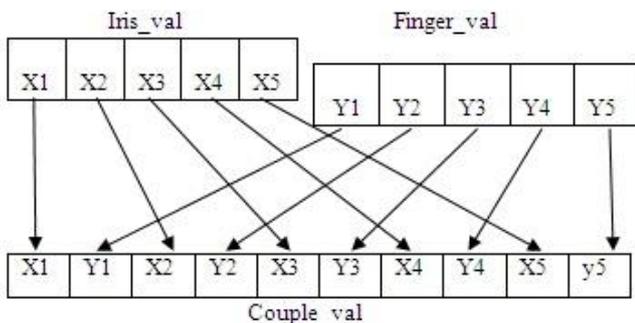


Figure 3. Generating the coupled value (Couple_val)

Then after creating the couple value apply the log to that value to obtain a high accuracy. This value can be called as the Authentication code or Identity code (Smart_Icode). The Smart Icode value is calculated as shown below.

$$\begin{aligned} \text{Smart_Icode} &= \log_{10}(\text{Couple_val}) \\ &= \log_{10}(1234567891) \\ &= 9.09151497752 \end{aligned}$$

This Smart Icode value is stored in the security database to provide authentication to the individuals to access the system.

IV. AUTHENTICATION PHASE

The authentication phase deals with providing authentication to the person. In this stage the individual's Iris and the Finger prints are scanned and the couple value

(Couple_val) is calculated and its log value (Smart_Icode) is computed. Now this Smart_Icode value is compared with value in the database.

4.1 GRANTING ACCESS

If the computed Smart_Icode value matches with the value stored in the database then the person is considered as the valid user and the access is granted for that person. Otherwise the access will not be provided to that person. The granting of access will be done as shown in Figure4.

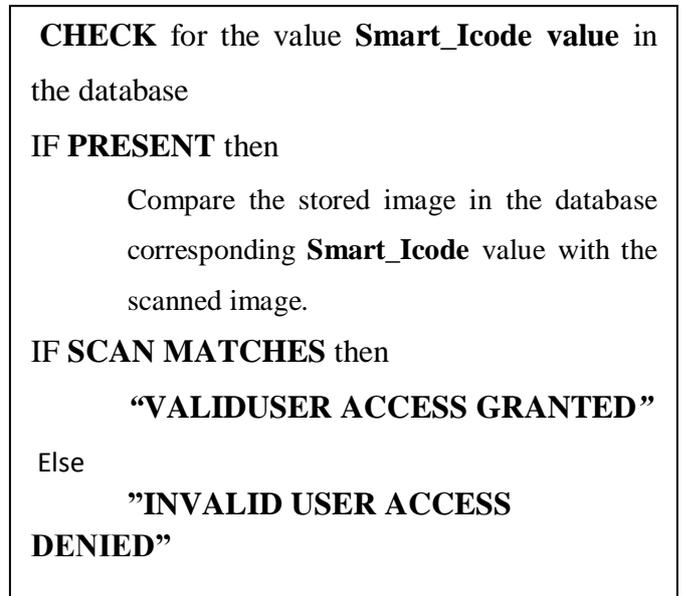


Figure 4. Authentication Process

V. ARCHITECTURE OF THE SECURITY SYSTEM

The Architecture of the security system works as shown in the figure 5. The architecture or the working of the Smart security System is clearly represented diagrammatically the working of the couple system is also displayed.

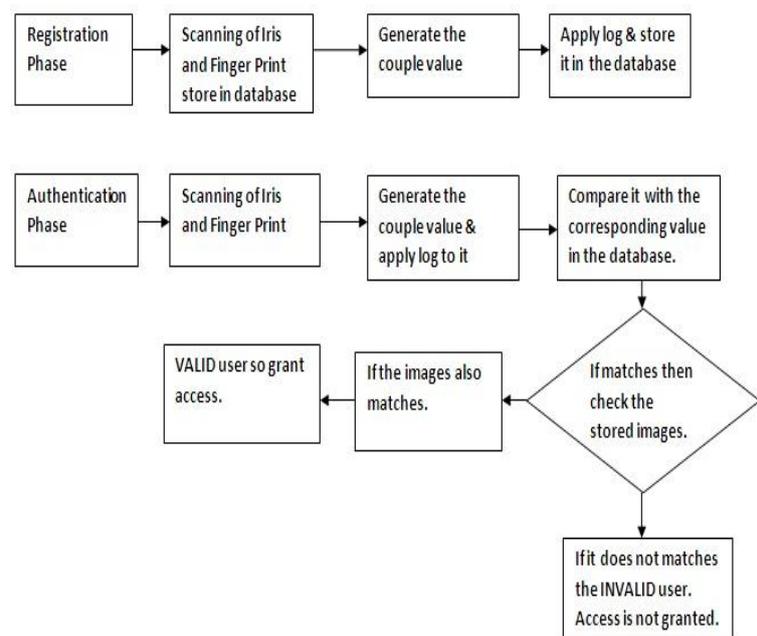


Figure 5. Architecture of the Smart coupled security system.

The working of the Registration and the authentication phase is separately displayed. The architecture show the flow in which the process is being carried out.

5.1 ADVANTAGES OF THIS SECURITY SYSTEM

The above design upon implementation would be much more advantageous and efficient than the present day designs that we use as a means that there is no other alternative way

- This new design is based on coupling the existing technologies rather trying our hands in a new sphere.
- The Hacker though possible of breaking into the first phase, his false identity would be revealed in our coupled design in the later part.
- There is no need to store FINGER PRINT IMAGES in the security database and hence efficient memory management.
- Comparison of numerical value is employed rather than image comparison hence the process is quickened and thus lot of time is saved
- The couple system value cannot be reversed and no intruders can understand the code even if they break it in the first stage.

VI. CONCLUSION

The Smart couple system of Biometrics is very efficient to provide a high level of security and authentication. The algorithm for the Smart couple system is very much easier to implement. This algorithm can be easily implemented using any programming language. Any future enhancement can be made to this system without much modifications or changes in the baseline. This is simple but efficient one and it uses the existing resources which is already available and known to us.

“You are the password No one else has it, but you”

As our future work we are going to include the Psycho-Physiology[10]. Psycho-physiological involves the measurement of physical symptoms, which reflect the emotional arousal of a person. For this purpose, we consulted with leading psychiatrists who confirmed us that the emotional state of a person will definitely be reflected in his physical behaviors, particularly during stress, feeling of guilt, tendency to cheat etc. Research done for a different purpose in various universities also confirmed that the physical characteristics such as heart activity, muscle activity, respiration, skin conductivity etc definitely reflect an individual’s mental state.

Acknowledgements

We would like to thank our institution for providing greater support in our Research.

REFERENCES

- [1] "Biometrics:Overview". Biometrics.cse. msu. edu. 2007-09-06.
- [2] Reference for eye biometrics <http://www.iris-scan.com/>
- [3] S. Dass and A. K. Jain, " Fingerprint Classification Using Orientation Field Flow Curves", *Proc. of Indian Conference on Computer Vision, Graphics and Image Processing*, (Kolkata), pp. 650-655, December 2004.

- [4] Y. Zhu, S. C. Dass, and Anil K. Jain, " Compound Stochastic Models for Fingerprint Individuality", *Proc. of International Conference on Pattern Recognition (ICPR)*, Vol. 3, pp. 532-535, Hong Kong, August, 2006.
- [5] S. C. Dass, Y. Zhu and Anil K. Jain, " Statistical models for assessing the individuality of fingerprints", *Fourth IEEE workshop on Automatic Identification Advanced Technologies*, pages 1-7,2005.
- [6] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the Individuality of Fingerprints", *IEEE Transactions on PAMI*, Vol. 24, No. 8, pp. 1010-1025, 2002. A shorter version also appears in *Fingerprint Whorld*, pp. 150-159, July 2002.
- [7] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM systems Journal*, vol. 40, pp. 614–634, 2001.
- [8] BIOMETRICS TO APPEAR Barry, S.C., Brooks, S.P., and Catchpole E.A. & Morgan B.J.T. (2002)
- [9] BIOMETRICS REDUNDANT MODEL Catchpole E.A., Morgan B.J.T. & Freeman S.N. (1998).
- [10] Hand book of Psycho-physiological second edition, edited by John T.Cacioppo,Louis G.Tassinary, Garry G.Berntson.

Experimental study of Roller burnishing process on plain carrier of planetary type gear box

P. S. Kamble¹, V. S. Jadhav²

¹P.G. Student of Govt.College of Engineering, Karad, 415124, Maharashtra, India.

²Professor of Mechanical Engg. Dept., Govt. College of Engineering, Karad, 415124, Maharashtra, India.

Abstract: Burnishing is cold working process where hard roller are being pressed against irregular surface, so that surface finish and the micro hardness increases. In this study internal roller burnishing tool is used to burnish the drilled hole. Speed, feed, and number of passes have been varied using taguchi method to examine the surface finish and microhardness. Anova analysis is carried out. Surface finish from 2.44 micron to 0.13micron is achieved.

Keywords: Roller burnishing, Surface roughness, Microhardness, drilling.

I. INTRODUCTION

Surface finish is required to avoid friction losses, good corrosion resistant property and high fatigue life. Conventional machining process leaves surface irregularities, which causes additional cost of finishing operations. Burnishing is a plastic deformation process. In the burnishing process, the pressure generated by the rollers exceeds the yield point of the softer piece part surface at the point of contact, resulting a small plastic deformation of the surface structure of the piece part. All machined surfaces consist of a series of peaks and valleys of irregular height and spacing. The plastic deformation created by roller burnishing displaces the material from the peaks by means of cold work under pressure into the valleys. The result is a mirror-like finish with a tough, work hardened, wear and corrosion resistant surface.[11]

Literature Review

A. Stoic et al.[1] investigated fine machining efficiency of 34CrMo4 steel using roller burnishing tool. Experimental results show that all smoothing outputs can be detected in all regimes. Roughness measured data before and after roller burnishing process have been compared. It was found that surface roughness is significantly lower after roller burnishing. Experimental results and numerical modeling of roller burnishing offers great potential in improving the efficiency and quality of parts. P. Ravindra Babu et al.[2] studied two internal roller burnishing tool to perform roller burnishing process on mild steel at different speed. The variation of surface finish and surface hardness are observed by varying speed.

Optimum increase in surface finish and surface hardness was at 62m/min. If speed is different than optimum value increase in surface finish is less. Basak et al.[3] carried out experiments using fuzzy model. Aluminium alloy (AL 7075 T6) has been burnished using different burnishing parameters such as number of revolution, feed, number of passes, and pressure force with burnishing apparatus. Using the experimental results a fuzzy logic model has been used to achieve the best parameters for the burnishing process. The fuzzy model prediction suggest that the most suitable values for minimum surface roughness are the pressure force

of 200 N, and a feed 0.1mm/rev with two tool passes. These results obtained from the fuzzy model are highly consistent with the experimental results except for a small deviation in the case of surface hardness value for 0.2 mm/rev of feed & 400N of applied force. The results shows that fuzzy logic is suitable technique that may be efficiently use to optimize burnishing process. Binu C Yeldose et al.[4] investigated comparison of effect of uncoated & Tin coated by reactive magnetron sputtering on EN31 rollers in burnishing with varying process parameters such as burnishing speed, feed, burnishing force, number of passes upon surface roughness of EN24 steel work material. It was observed that the performance of the Tin-coated roller is superior to uncoated rollers in burnishing operation. The burnishing speed, feed, depth of cut and number of passes are influencing parameters on the burnishing operation. The burnishing speed, burnishing force and number of passes are having almost equal importance on the performance of the roller in burnishing, particularly with reference to the surface finish of the components produced. S. Thamizhmani et al.[5] investigated surface roughness and surface hardness by burnishing on titanium alloy. The test results produce improvement in surface finish. Studies can be extended to fatigue testing after burnishing process. T. Altan et al.[6] developed 2D & 3D FEM model was used to study the effect of process parameters burnishing pressure & feed rate on surface finish and residual stresses. The simulation results were evaluated and compared with the experimental data. Results shows that the established FEM model could predict the residual stresses and provide useful information for the effect of process parameters. Both FEM and experiments shows that burnishing pressure is the most influence, where high burnishing pressure produces less roughness and more compressive residual stress at the surface. S.Hassan et al.[7] studied multi roller burnishing on non ferrous materials, namely aluminium, brass and copper to improve surface finish and surface hardness. The surface roughness on non ferrous metal is improved by high feed rate and depth of penetration. N.S.M. El-Tayeb et al.[8] developed a simple burnishing tool, with interchangeable adapter for ball & roller. They fabricated the tool to perform roller burnishing processes on Aluminium 6061 under different parameters & different burnishing orientation. The impact of burnishing speed, force and burnishing tool dimensions on the surface qualities and tribological properties were investigated. It was found that burnishing speed of 330 rpm and burnishing force of 212N is capable of improving surface roughness as much as 40%. However, an increase in the roller contact width leads to less improvement in the surface roughness. In tribo test, burnished surface using small roller produces the lowest friction coefficient under dry contact condition. A substantial decrease in the friction coefficient and weight loss was obtained under lubricated contact condition. Furthermore, a 46% reduction in the friction coefficient was

obtained when sliding took place in the direction parallel to burnishing orientation. These findings are further supplemented by SEM photomicrographs of tested surfaces. H. Hamadache et al.[9] developed a device for mechanical plastic deformation of structural Rb40 steel using ball and roller burnishing and investigated the change in roughness, hardness and wear resistance. They found that roller burnishing provide optimal roughness results, particularly when initial surface quality is close to 3 μm. In terms of hardness, ball burnishing process becomes interesting. For both cases, the optimum roughness and hardness are obtained for a specific regime whose decisive parameters are the applied force as well as number of burnishing tool passes. Based on roughness, it is recommended to limit the number of passes to two where as for the highest hardness, it is advised to go up to three passes while associating an effort of 150N. B.B. Ahuja et al.[10] carried out experiments based on 2³ factorial designs on turn master T-40 lathe. They studied that the effect of the combined turning and two ball burnishing parameters on the surface roughness and surface hardness of aluminum specimen. The results have been analyzed by the variance technique and the F-test. Analysis shows that the lubricant, force, speed and feed have significant effects on surface roughness and surface hardness.

II. MATERIAL AND METHOD:

In this current research paper, an effort is being made to understand the improvement in the surface finish and microhardness of burnished surfaces along with the influence of the process parameters in “En-8” material, which is used as plain carrier in planetary type gear box to carry the planetary gears.

TABLE 1.Composition of EN-8

%C	%Mn	%Si	%S	%P
0.38	0.65	0.23	0.050max	0.050max



Fig.1 Specimen (Plain carrier) of EN-8 material

Fig.1 shows the specimen prepared for this study. The work piece is of En-8 material has six number holes each. These holes are drilled to 16mm diameter on drilling machine. Burnishing is carried on DRO type boring machine shown in fig.2. Internal roller burnishing tool (fig.3) has five number of roller on the periphery, which pressed against the peaks on inside surface of the hole.

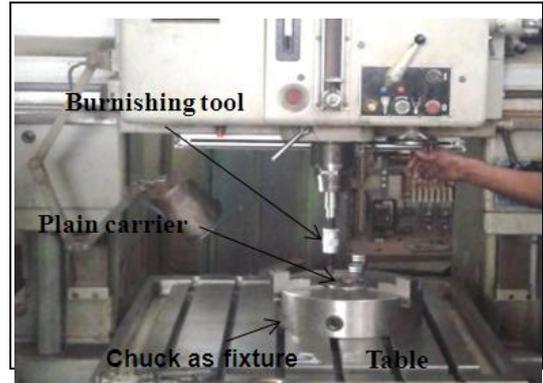


Fig.2 Experimental set up (Top gear transmission, satara)



Fig.3 Roller burnishing tool

TABLE 2.Process parameter and their range.

levels	1	2	3
Parameters			
Speed(rpm)	560	800	1125
Feed(mm/rev)	0.04	0.12	0.21
No.of passes	1	2	3

In this study Taguchi L9 Orthogonal array has been used to check the results of input parameters. Process parameters are shown in table.2. Surface roughness values were measured before burnishing and after burnishing of each trial using roughness measurement device (Hommelwerke).

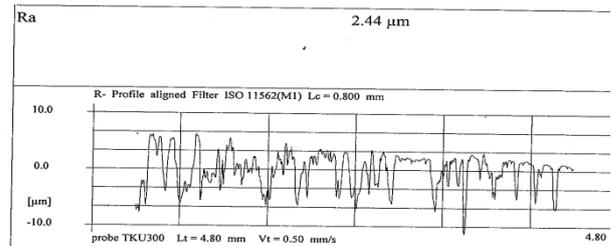


Fig.4 Roughness before burnishing

Comparison has been made with the help of graph plotted by roughness tester shown in fig.4&5. There is significant change in plastic deformation of higher picks after burnishing.

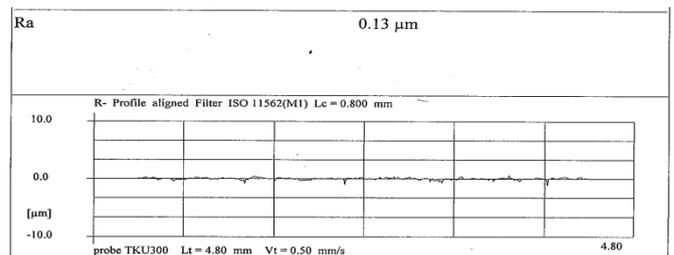


Fig.5 Roughness after burnishing

III. Results & Discussion

From the trials it is found that there is significant change in surface roughness due to the variation in spindle speed, feed and number of passes.

IV. Effect of feed:

Maximum surface finish is achieved at higher feed 0.21mm/rev shown in fig.6. As feed increases roughness is decreases.

TABLE 3. L9 Taguchi Orthogonal array & observation

Experiment no.	Speed (rpm)	Feed (mm/rev)	No. of passes	Roughness (micron)	S/N
E1	560	0.04	1	0.63	4.013
E2	560	0.12	2	1.21	-1.655
E3	560	0.21	3	1	0
E4	800	0.04	2	0.17	15.391
E5	800	0.12	3	0.23	12.765
E6	800	0.21	1	0.13	17.721
E7	1125	0.04	3	0.9	0.915
E8	1125	0.12	1	0.19	14.424
E9	1125	0.21	2	0.24	12.395

TABLE 4. Response table of S/N ratio for smaller is better

Level	Speed	Feed	No. of passes
1	0.7858	6.7731	12.0531
2	15.2925	8.5116	8.7104
3	9.2453	10.0390	4.5602
Delta	14.5067	3.2658	7.4929
Rank	1	3	2

TABLE 5. Analysis of variance

Source	D.O.F.	Seq. SS	Adj. SS	Adj. Mean Sq.	F	P
Speed	2	318.58	318.58	159.288	9.14	0.099
Feed	2	16.02	16.02	8.010	0.46	0.685
No. of passes	2	84.54	84.54	42.270	2.42	0.292
Residual error	2	34.87	34.87	17.436	*	*
Total	8	457.01	*	*	*	*

SS = Sum of Squares, D.O.F. = Degree of Freedom, SS = Sum of squares.

TABLE 6. Optimal combination for better surface finish

Parameter	Speed	Feed	No. of passes
Better surface finish	800	0.21	1

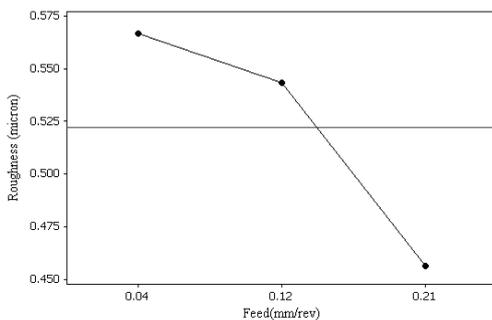


Fig.6 Feed Vs Surface roughness

Effect of speed:

Maximum surface finish is achieved at 800rpm. It is found that as speed increases than optimum value roughness start increasing shown in fig.7.

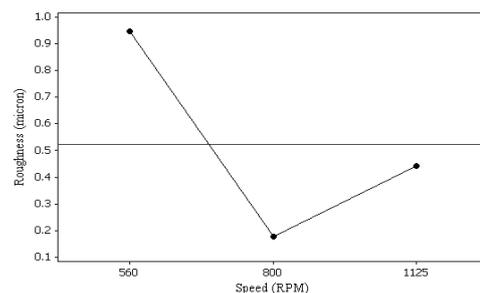


Fig.7 Spindle revolution Vs Roughness

Effect of number of passes:

Number of passes is also important factor. It is found that at one number of pass burnishing gives better microhardness. As number of passes increases roughness start increasing shown in fig.8.

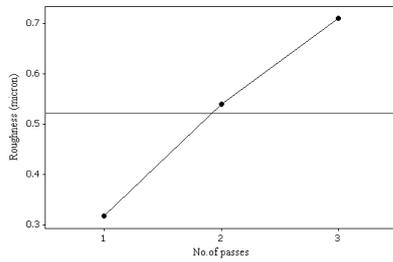


Fig.8 No. of passes Vs Roughness

The ANOVA results and optimal combination for surface finish are shown in Tables.5 & 6 respectively.

Study of micro hardness:

In this study microhardness testing is carried out on inside surface of burnished hole using Wilson instrument/402MVO microvicker hardness tester. Table.7. shows the microhardness reading taken randomly on burnished surface. Microhardness testing images of before and after burnishing are shown in fig.9.

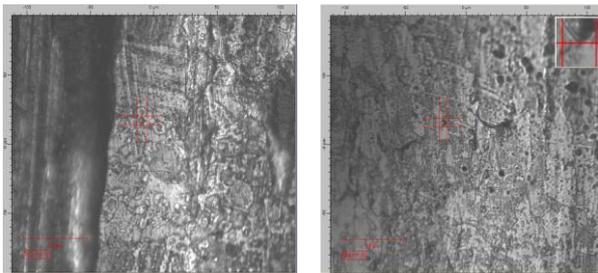


Fig.9 Microhardness testing image by tester(a)Before burnishing(b)After burnishing

Effect of feed:

Maximum microhardness is achieved at lower feed. At 0.04mm/rev feed value microhardness is achieved maximum shown in fig.10.

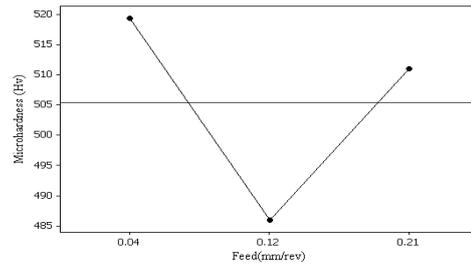


Fig.10 Microhardness Vs Feed

Effect of speed:

Microhardness increases as speed increases. It is observed that at 1125rpm speed microhardness value is maximum shown in fig.11.

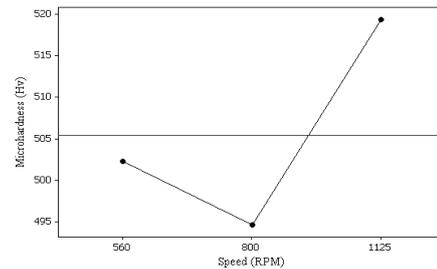


Fig.11 Microhardness Vs Speed

Effect of number of passes:

It is found that at one number of pass, microhardness is maximum. As number of passes increases microhardness decreases shown in fig.12.

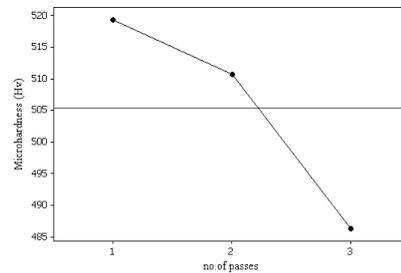


Fig.12 Microhardness Vs No. of passes

The ANOVA results and optimal combination for Microhardness are shown in Tables.9 & 10 respectively.

TABLE.7. L9 Taguchi Orthogonal array & observation

Experiment no.	Speed (rpm)	Feed (mm/rev)	No. of passes	Microhardness (Hv)	S/N
E1	560	0.04	1	528	54.452
E2	560	0.12	2	502	54.014
E3	560	0.21	3	477	53.570
E4	800	0.04	2	502	54.014
E5	800	0.12	3	454	53.141
E6	800	0.21	1	528	54.452
E7	1125	0.04	3	528	54.452
E8	1125	0.12	1	502	54.014
E9	1125	0.21	2	528	54.452

TABLE.8. Response table of S/N ratio for larger is better

Level	Speed	Feed	No. of passes
1	54.01	54.31	54.31
2	53.87	53.72	54.16
3	54.31	54.16	53.72
Delta	0.44	0.58	0.59
Rank	3	2	1

TABLE 9. Analysis of variance

Source	D.O.F.	Seq.S.S.	Adj. S.Sq.	Adj. Mean Sq.	F	P
Speed	2	0.2981	0.2981	0.1491	0.99	0.502
Feed	2	0.5519	0.5519	0.2759	1.83	0.353
No. of passes	2	0.5563	0.5563	0.2782	1.85	0.351
Residual error	2	0.3008	0.3008	0.1504	*	*
Total	8	1.7071	*	*	*	*

SS = Sum of Squares, D.O.F. = Degree of Freedom, SS = sum of squares.

TABLE 10. Optimal combination for better surface finish

Parameter	Speed	Feed	No.of passes
Better surface finish	1125	0.04	1

V. Conclusions

1. Roller burnishing produce superior finish. Ra value observed is finest upto 0.13 micron (Table.3).
2. Before burnishing microhardness found 377Hv and after burnishing it increases upto 528Hv (fig.9).
3. Many researchers carried out experiments on external work piece by single roller burnishing tool. From this study it is observed that roller burnishing also gives better result in drilled hole.

Acknowledgement

The authors gratefully acknowledge Top Gear Transmission, Satara for the cooperation and help. The authors are also thankful to UGC Lab division, Pune and TMC measurement lab, Pune for providing the microhardness and roughness testing facilities respectively.

REFERENCES

- [1] A. Stoić, I. Lackovic, J. Kopac, I. Samardzic, D. Kozak, "An Investigation of Machining Efficiency of Internal Roller Burnishing", Journal of Achievements in Material and Manufacturing Engineering, Volume 40, Issue2, June 2010.
- [2] P. Ravindra babu, T. Siva Prasad, A. V. S. Raju, A. Jawahar Babu, "Effect of Internal Roller Burnishing on Surface Roughness and Surface Hardness of Mild Steel", Journal of Scientific & Industrial Research, Volume.68, January 2009, pp. 29-31.
- [3] Hudayim Basak, H. Haldun Goktas, "Burnishing Process on Al-alloy & Optimization of Surface Roughness and Surface Hardness by Fuzzy Logic", Journal of Material & Design, 30(2009) 1275-1281.
- [4] Binu C. Yeldose, B. Ramamoorthy, "An Investigation into the High Performance of Tin-coated Rollers in Burnishing Process", Journal of material processing technology, 207 (2008) 350-355.
- [5] S. Thamizhmani, B. Bin Omar, S. Saporudin, S. Hasan, "Surface Roughness Investigation and Hardness by Burnishing on Titanium Alloy", Journal of Achievements in Material and Manufacturing Engineering, Volume 28, Issue2, June 2008
- [6] Taylan Altan, Partchapol Sartkulvanich, Francisco Jasso, Ciro Rodriguez, "Finite Element Modeling of Hard Roller Burnishing: An Analysis on the effect of Process Parameters upon Surface Finish and Residual stresses", ASME Journal of Manufacturing Science and Engineering, August 2007, Vol. 129
- [7] S. Hassan, S.Thamizhmanii, B. Saporudin, S. Hasan, "A Study of Multi-Roller Burnishing on Non-Ferrous Metals", Journal of Achievements in Material and Manufacturing Engineering, Volume 22, Issue2, June 2007.
- [8] N.S.M.El-Tayeb, K.O.Low, P.V.Brevem, "Influence of Roller Burnishing Contact Width and Burnishing Orientation on Surface Quality and Tribological Behavior of Aluminium 6061", Journal of Material Processing Technology 186 (2007) 272-278.
- [9] H.Hamadache, L.Laouar, N.E.Zeghib, K.Chaoui, "Characteristics of Rb40 Steel Superficial Layer under Ball and Roller Burnishing", Journal of Material Processing Technology 180(2006) 130-136.
- [10] B. B. Ahuja & U. M. Shirsat, "Parametric Analysis of Combined Turning and Ball Burnishing Process", Indian Journal of Engineering & Material Science, Vol.11, October 2004, pp.391-396.
- [11] P.S.Kamble, V.S.Jadhav & C.Y.Seemikeri, "Surface finish improvement by Roller Burnishing Process on En-24", Proceeding of National Conference on Recent Advancements in Engineering-2012, India. Vol.1, pp.197-200. ISBN:978-93-82062-37-0.

Effect of Cracked Section on Lateral Response of Reinforced Concrete Flanged Beams

Wakchaure M. R.¹, Varpe Charulata S.²

1(Asso.Professor, Civil Engg. Dept., Amrutvahini College of Engineering, Maharashtra, India)

2(Assi.Professor, Civil Engg. Dept., MIT College of Engineering, Pune, Maharashtra, India)

Abstract: In an analysis of reinforced concrete structure the flexural stiffness of section is an important parameter and the change in the value of flexural stiffness may result in significant change in analysis result. The current study aims to estimate the reduction in flexural stiffness of reinforced concrete flanged beam sections subjected to lateral loading and take the cracking effect of reinforced concrete section into account. The reduction in flexural stiffness due to earthquake shaking may increase the lateral deflection and it can be significantly greater as compared to deflection estimated using gross flexural stiffness. To take this effect into consideration, the design code of some countries suggest reduction factors or equations to reduce the gross flexural stiffness to effective flexural stiffness but they have some drawbacks because not consider all important parameters in there equation. However in Indian seismic code IS 1893 (2002) and many countries there are no provisions to account for reduction in stiffness due to concrete cracking. Analytical work in present study identified the parameters and influences of these parameters on effective stiffness are determined and suggest the simplified but reasonable accurate expression for computation of effective stiffness of reinforced concrete flanged beam section. Proposed equations can be easily use by designer to estimate the effective stiffness of reinforced concrete flanged beam sections accurately.

Keywords: Concrete cracking, Deflections, effective stiffness, flanged beam, regression analysis.

I. Introduction

While analyzing the reinforced concrete structures under vertical and lateral loads, the designers are consider the assumed value of the flexural stiffness, but under the combined action of vertical and lateral loads, some sections within critical member will reach near yield point resulting in cracking of the member on bending tensile side, the flexural stiffness (EI) of member starts decreasing or reducing. There will be considerable reduction in flexural stiffness due to cracking. Because of reduction in flexural stiffness value, the lateral deflection of reinforced concrete members increases and it can be significantly greater as compared to deflection estimated using gross flexural stiffness. Also the natural time period, deflection, internal force distribution and dynamic response all changes due to change in stiffness (EI) i.e. whole analysis changes therefore it is essential to use the reduced or effective stiffness of reinforced concrete structure. To take these effects into consideration the design code of many countries suggests some reduction factor or equations to reduce the gross stiffness to effective stiffness. However in Indian seismic code IS 1893 (2002) and many countries there are

no provisions to account for reduction in stiffness due to concrete cracking.

While analyzing the reinforced concrete structures under vertical and lateral loads, the designers are consider the assumed value of the flexural stiffness, but under the combined action of vertical and lateral loads, some sections within critical member will reach near yield point resulting in cracking of the member on bending tensile side, the flexural stiffness (EI) of member starts decreasing or reducing. There will be considerable reduction in flexural stiffness due to cracking. Because of reduction in flexural stiffness value, the lateral deflection of reinforced concrete members increases and it can be significantly greater as compared to deflection estimated using gross flexural stiffness. Also the natural time period, deflection, internal force distribution and dynamic response all changes due to change in stiffness (EI) i.e. whole analysis changes therefore it is essential to use the reduced or effective stiffness of reinforced concrete structure. To take these effects into consideration the design code of many countries suggests some reduction factor or equations to reduce the gross stiffness to effective stiffness. However in Indian seismic code IS 1893 (2002) and many countries there are no provisions to account for reduction in stiffness due to concrete cracking.

II. Methodology

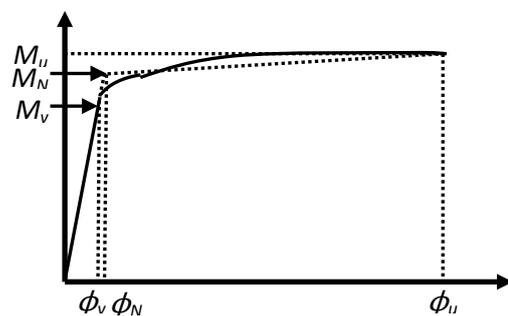


Fig. 1: Typical moment curvature relationship and its bilinear approximation.

To find out the reduction in stiffness of reinforced concrete section due to concrete cracking, beam sections have been analyzed. Moment curvature relationships have been obtained by using confined Mander's stress strain curve for concrete and simple stress strain curve for steel reinforcement for different sections. Figure 1 show typical moment curvature relationship and its bilinear approximation. It is a well known fact that the most approximate linearization of moment curvature relation is by an initial elastic segment passing through first yield and extrapolated to the nominal flexural strength M_N , and a post yield segment connected to the ultimate strength. Then the flexural stiffness at yield level (EI_y) has been calculated

using Eqn. (1), when section first attains the reinforcement tensile yield strain of $\epsilon_y = \frac{f_y}{E_s}$, or concrete extreme compression fiber attains a strain of 0.002, whichever occurs fir

$$EI_y = \frac{M_y}{\phi_y} \tag{1}$$

Where, EI_y is flexural stiffness at yield level, M_y is moment capacity at yield level and ϕ_y is yield curvature. The nominal flexural strength M_N develops when the extreme compression fiber strain reaches 0.004 or the reinforcement tension strain reaches 0.015, whichever occurs first (Priestley 2003). So reduce or effective flexural stiffness at this level is given by Eqn. (2)

$$EI_{eff} = \frac{M_N}{\phi_N} \tag{2}$$

Where, EI_{eff} is the reduced flexural stiffness, M_N is the nominal flexural strength and ϕ_N is the curvature corresponding to M_N .

Finally EI_{eff} and EI_y are normalized with EI_{gross} i.e. flexural stiffness of gross cross section, where $E = 5000\sqrt{f_{ck}}$ as per IS 456 (2000) and $I_{gross} = bD^3/12$. Here f_{ck} is the characteristic compressive strength of concrete cubes of size 150 mm at 28 days. The reinforced concrete flanged beam sections varying in b_w/b_f ratio from 0.135 to 0.255 and in D_f/D ratio from 0.192 to 0.277 were analyzed for percentage of steel varying from 0.4 to 2.5. Finally all these flanged beam sections were analyzed for two different concrete grade having f_{ck} equal to 20 Mpa and 25 Mpa) and reinforcement grade having f_y equal to 415 Mpa and 500 MPa .To know the variation of EI_{eff} of beam section with one parameter all other parameters were kept constant, e.g., to know the variation of EI_{eff} with percentage of steel, the aspect ratio, grade of concrete and reinforcement, confinement reinforcement were kept same and only percentage of steel was changed.

2.1 Typical Moment Curvature Relationship

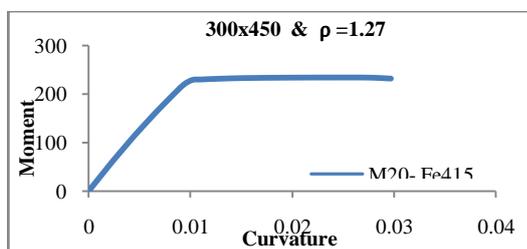


Fig. 2: Typical moment curvature relationship

Approximately 400 flanged beam sections were analyzed in section designer of computer program SAP 2000 V14 (2009). Moment curvature relationships have been obtained in section designer by using confined Mender’s stress strain curve for concrete and simple stress strain curve for steel reinforcement for different sections. Then the EI_{eff} was found out by using Eqn. (2). Figure 2 shows typical moment curvature relationship obtained from section designer for rectangular beam section.

2.2 Influence of Different Type of Stress-Strain Curve of Concrete on Young’s Modulus (E)

Different researchers have used the different stress strain curve of concrete for developing the moment curvature relationship for reinforced concrete sections. In

the current study confined Mander’s stress strain curve for concrete is used. To notice the difference between E values in different stress strain curves for concrete, Figure 3 is plotted.

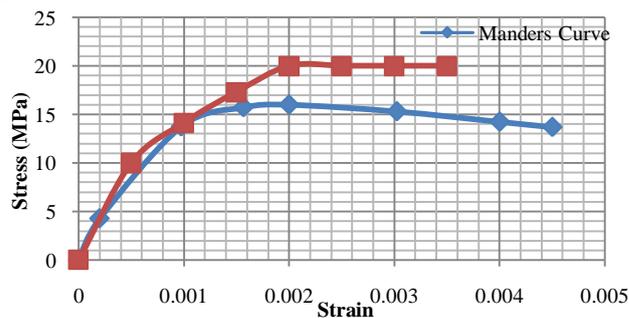


Fig. 3: Stress strain curves for concrete under compression

Figure 3 shows stress strain curve for M25 grade of concrete proposed by Mander and parabolic stress strain curve given in IS 456 (2000). From above figure it was clear that the difference in value of E is very insignificant up to strain level of 0.0015, but on further increase in strain level of concrete the difference in E value increased significantly. The value of E obtained by using parabolic stress strain curve was approximately 1.33 and 1.5 times the value obtained by using Mander’s stress strain curve, when strain value in concrete was 0.003 and 0.004, respectively. In the current study equations were proposed to estimate the effective stiffness (EI_{eff}) value of reinforced concrete rectangular and flanged beam sections by using Mander’s stress strain curve [12]. To obtain the effective stiffness (EI_{eff}) value corresponding to parabolic or design stress strain curve it is recommended that effective stiffness (EI_{eff}) value should be multiplied by factor of 1.33 when strain in concrete is 0.003 and by a factor of 1.5 when strain in concrete is 0.004, respectively.

III. Analysis of Reinforced Concrete (RC) Flanged Beam Sections

There were up to 400 flanged beam sections analyzed here. In this analysis case observed here is the estimation of $EI_{eff}/EI_{gross (Rect.)}$ i.e. in these case the gross stiffness calculate for rectangular section of flanged beam. In each case ten different reinforced concrete flanged beam section varying with different parameter such as percentage of steel, b_w/b_f and D_f/D ratio of section and grade of concrete and steel. out of this, in five different flanged beam section keep D_f/D ratio constant and b_w/b_f ratio varying from 0.135 to 0.255 and in other five different flanged beam section keep b_w/b_f ratio constant and D_f/D ratio varying from 0.192 to 0.277. Figure 4 show flanged beam section.

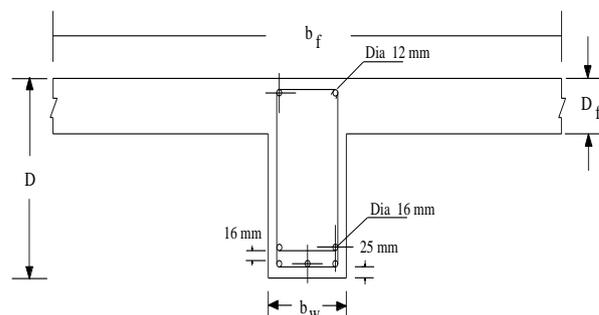


Fig. 4: Flanged beam diagram

3.1 Effective Stiffness of Reinforced Concrete (RC)

Flanged Beam Sections (EI_{eff}/EI_{gross} (Rect.))

Determine the effective stiffness of reinforced concrete flanged beam section, in these case the gross moment of inertia ($b_w D^3/12$) was calculated for rectangle section of flanged beam. EI_{eff}/EI_{gross} (Rect.) is varying with the parameter same like above such as percentage of steel, D_f/D ratio, b_w/b_f ratio, grade of concrete and steel.

3.1.1 Keeping D_f/D Constant and b_w/b_f Varying

Keep D_f/D constant and b_w/b_f varying from 0.135 to 0.255

a) Variation of EI with Percentage of Steel

In the current study the percentage of steel considered in bottom of section was from nearly 0.4 (just above the minimum) to maximum of up to 2.5 percent. The minimum of two bars were provided at top, because of need for provision of confinement reinforcement (2 legs). The section was developed in section designer of SAP2000 V14. Figure 5 shows variation of EI_{eff}/EI_{gross} of reinforced concrete flanged beam sections with percentage of steel. EI_{eff}/EI_{gross} is directly proportional to percentage of steel

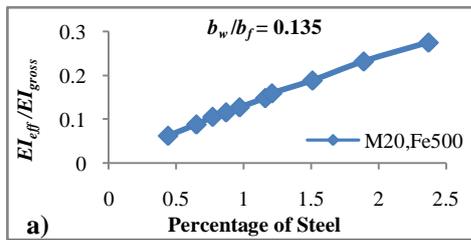


Fig. 5: Variation of EI with percentage of steel for different b_w/b_f ratio

b) Variation of EI with Grade of Concrete and Steel

From Figure 6 it may be observed that EI_{eff}/EI_{gross} of flanged beam sections increases with increase in grade of steel and decreases with increase in concrete grade steel for b_w/b_f ratio 0.135 and 0.255 it also same for b_w/b_f ratio 0.153, 0.177 and 0.209. Hence it may be said that EI_{eff}/EI_{gross} of flanged beam sections was directly proportional to grade of steel, while it was inversely proportional to grade of concrete.

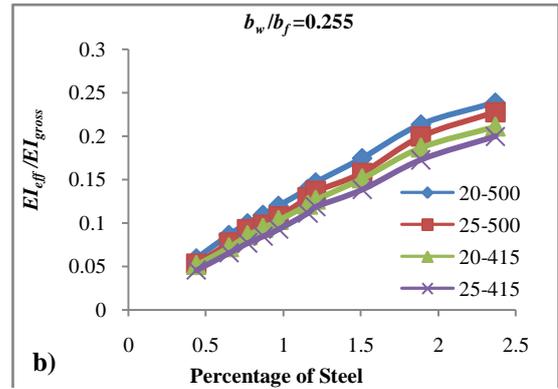
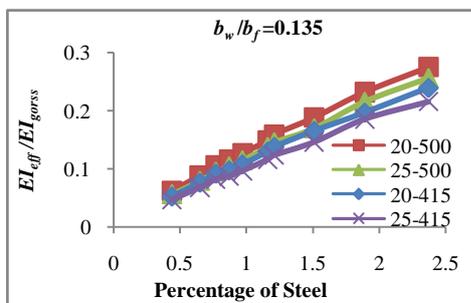


Fig. 6: Variation of EI with grade of concrete and steel for different b_w/b_f ratio

c) Variation of EI with b_w/b_f Ratio of Reinforced Concrete (RC) Flanged Beam Section

From Figure 7 it may be clearly observed that EI_{eff}/EI_{gross} of flanged beam sections increases with decrease in b_w/b_f ratio for same percentage of steel in section. i.e.; EI_{eff}/EI_{gross} of flanged beam was inversely proportional to b_w/b_f ratio. This observation was valid for percentage of steel ranging from 1 to 2.5. The EI_{eff}/EI_{gross} of flanged beam sections for very low level of steel percentage (upto 1) was almost constant.

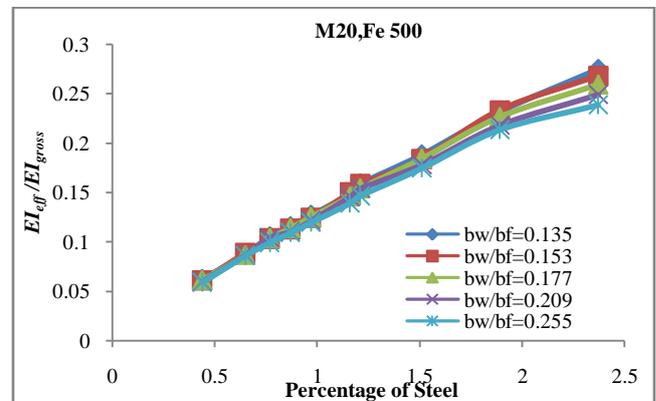
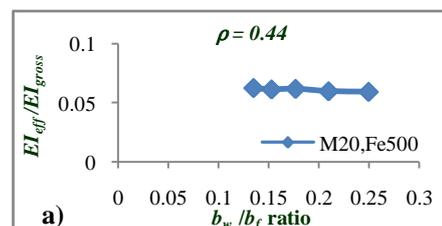


Fig. 7: Variation of EI with b_w/b_f ratio

d) Variation of EI with b_w/b_f Ratio for same Percentage of Steel

From Figure 8 it clearly observed that EI_{eff}/EI_{gross} of flanged beam sections is almost constant in case of percentage of steel 0.44 for different b_w/b_f ratio. And EI_{eff}/EI_{gross} increases with decrease in b_w/b_f ratio having percentage of steel 2.37, the difference in EI_{eff}/EI_{gross} due to same percentage of steel 2.37 is relatively significant.



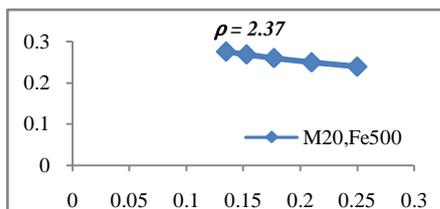


Fig. 8: Variation of EI with b_w/b_f ratio for same Percentage of Steel

3.1.2 Keep b_w/b_f Constant and D_f/D Varying

Keep b_w/b_f constant and D_f/D varying from 0.192 to 0.277

a) Variation of EI with Percentage of Steel

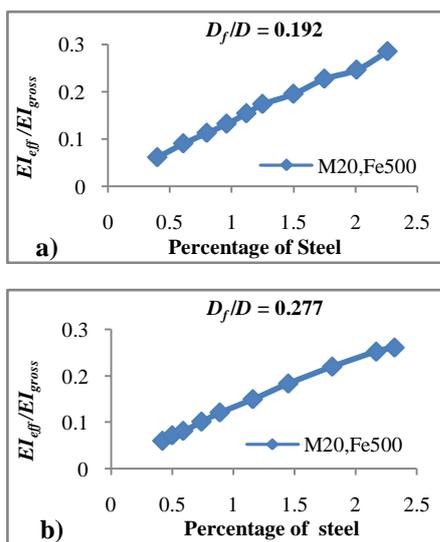


Fig. 9: Variation of EI with percentage of steel for different D_f/D ratio

From Figure 9, it may be observed that the EI_{eff}/EI_{gross} of flanged beam sections increases with increase in percentage of steel for b_w/b_f ratio 0.192 and 0.277 i.e.; EI_{eff}/EI_{gross} is directly proportional to percentage of steel.

b) Variation of EI with Grade of Concrete and Steel

From Figure 10, observed that EI_{eff}/EI_{gross} of flanged beam sections was directly proportional to grade of steel, while it was inversely proportional to grade of concrete.

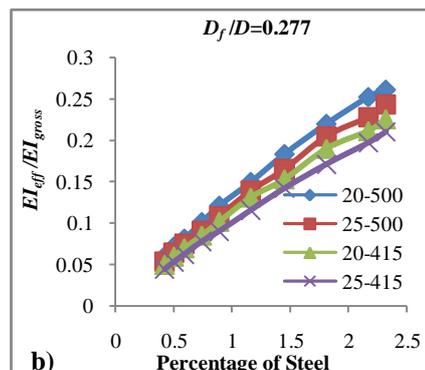
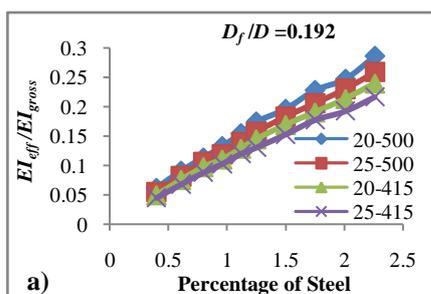


Fig. 10: Variation of EI with grade of concrete and steel for different D_f/D ratio

c) Variation of EI with D_f/D Ratio of RC Flanged Beam Section

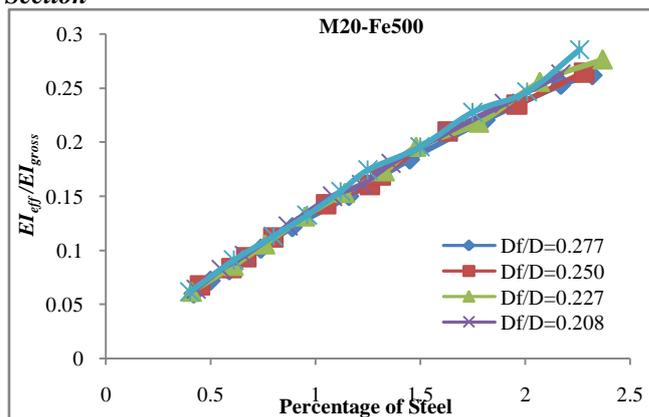
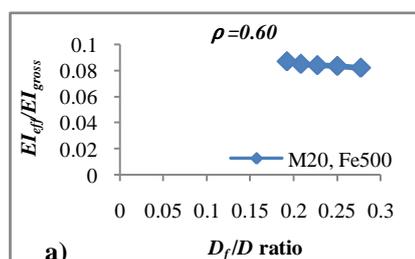


Fig. 11: Variation of EI with D_f/D ratio

From Figure 11, observed that EI_{eff}/EI_{gross} of flanged beam sections increases with decrease in D_f/D ratio for same percentage of steel in section. i.e.; EI_{eff}/EI_{gross} of flanged beam was inversely proportional to D_f/D ratio. This observation was valid for percentage of steel ranging from 1 to 2.5. The EI_{eff}/EI_{gross} of flanged beam sections for very low level of steel percentage (upto 1) was almost constant.

d) Variation of EI with D_f/D Ratio for same Percentage of Steel

From Figure 12 it observed that EI_{eff}/EI_{gross} of RC flanged beam sections is almost constant in percentage of steel 0.6 for different D_f/D ratio. And EI_{eff}/EI_{gross} increases with decrease in D_f/D ratio having percentage of steel above the 1 percent



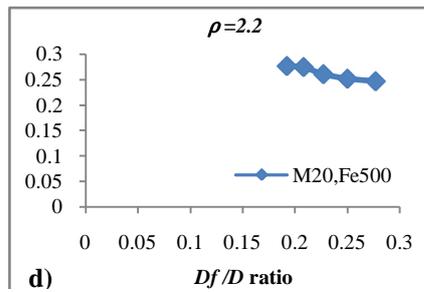


Fig. 12: Variation of EI with D_f/D ratio for same percentage of steel

3.1.3 Effective stiffness of RC flanged beam section ($EI_{eff}/EI_{gross(Rect.)}$)

Linear unconstrained regression analysis was carried out on data obtained from analysis of nearly 400 reinforced concrete flanged beam sections with different section parameters. The Eqn. (3) was proposed to estimate the EI_{eff}/EI_{gross} of RC flanged beam sections. The equation estimates the EI_{eff}/EI_{gross} at concrete strain of 0.004 in extreme compression fiber.

$$\frac{EI_{eff}}{EI_{gross(Rect.)}} = 0.08 \left(\frac{f_y^{0.87} \rho_B^{0.88}}{f_{ck}^{0.42} (D_f/D)^{0.18} (b_w/b_f)^{0.09}} \right) \quad (3)$$

Where, f_{ck} = grade of concrete in MPa, D_f/D = ratio of flanged depth to total depth of section, b_w/b_f = ratio of web width to flanged width, ρ_B = Steel area at bottom expressed as fraction of cross sectional area of section (in decimal), $I_{gross(Rect.)}$ = gross moment of inertia of rectangle section in flanged beam.

Coefficient of correlation between analytically obtained values of EI_{eff}/EI_{gross} and values obtained using developed equations for flanged beam sections was found out to be 0.99. If parabolic or design stress strain curve is used it is recommended that to account for the difference in Young's Modulus (E) with different stress strain curve the value of EI_{eff}/EI_{gross} obtained by Eqn. (3) should be multiplied by factor of 1.5 if strain in concrete at extreme fiber is 0.004.

$$\left[\frac{EI_{eff}}{EI_{gross(Rect.)}} \right]_{(P.S.S)} = 1.5 \left[\frac{EI_{eff}}{EI_{gross(Rect.)}} \right]_{(M.S.S)} \quad (4)$$

At concrete strain of 0.004

Where, P.S.S = Parabolic or design stress strain curve and M.S.S = Stress strain curve proposed by Mander

IV. Verification of Proposed Equation

To verify the proposed equations, the values obtained from proposed equations were compared with the equations developed by other researchers and from experimental data published in various literatures. The equations developed

for estimating effective stiffness of flanged beams were verified by comparison with equations or values proposed by khuntia and Ghosh (2004) and Priestley (2003). It was noticed that the equations proposed by other researchers were holding good only up to particular limits, also some of important parameters were not considered by other researchers in there equations and not estimating the value of effective stiffness of reinforced concrete flanged beam sections accurately. Following are equations proposed by Khuntia and Ghosh (2004) for effective stiffness determination of reinforced concrete rectangular and flanged beam sections.

For rectangular beam sections-

$$EI_{eff} = E_c I_g (0.10 + 25\rho) \left(1 - 0.2 \frac{b}{d} \right) \leq 0.6 E_c I_g \quad (5)$$

For flanged beam sections-

$$\frac{EI_{eff(T)}}{EI_{eff}} \left(1 + 2 \frac{t_f}{h} \right) \leq 1.4 \quad (6)$$

Where: b/d = aspect ratio or ratio of width to depth of beam; ρ = steel area express as fraction of cross sectional area of section (in decimal); t_f/h = ratio of flanged thickness to total depth. $EI_{eff(T)}$ = effective stiffness of flanged beam section. EI_{eff} = effective stiffness of rectangular beam section from Eqn. (5)

Proposed Eqn. (3) in terms of Parabolic or design stress strain curve is that Eqn. (4) were compared with values given by Priestley (2003) and values obtained from Eqn. (6) proposed by Khuntia and Ghosh (2004). Table 1, shows the values of effective stiffness estimated by using Khuntia and Ghosh Eqn. (6), proposed Eqn. in current study Eqn. (3) and values proposed by Priestley (2003).

It may be clearly observed from Table 1 that the difference in values of estimated effective stiffness by using Eqn. (6) proposed by Khuntia and Ghosh (2004) and that proposed in current study was significant; though the difference in values of estimated effective stiffness by using Eqn. (3) proposed in current study and that proposed by Priestley (2003) was also significant. The reason for difference was some limitations of equation proposed by Khuntia and Ghosh (2004) and Priestley (2003), which are discussed later in next point.

4.1 Limitation of Existing Equations

In case of RC flanged beam sections, From Khuntia & Ghosh Eqn. (6) it observed that there was no parameter to incorporate the effect of b_w/b_f ratio and grade of reinforcement which are important parameter in effective stiffness determination of flanged beam sections. Because of this reason the difference estimated between Khuntia & Ghosh (2004) Eqn. value and that obtained from Eqn. (3)

Table I. Comparison of EI_{eff} values of flanged beam section proposed by different researcher

b_w (mm)	b_f (mm)	D_f (mm)	D (mm)	f_{ck} MPa	f_y MPa	% ρ_B Bottom	Ghosh Eqn. (6)	Priestley Eqn. (2003)	Prop. Eqn. (3)	Prop. Eqn. x 1.5
350	1550	150	600	30	400	0.82	0.17	0.17	0.076	0.11
350	1550	150	600	30	400	1.54	0.27	0.36	0.132	0.20
350	1550	150	600	30	400	2.2	0.37	0.49	0.180	0.27
350	1550	150	600	30	400	2.2	0.37	0.27	0.180	0.27
350	1550	150	600	30	300	0.82	0.17	0.19	0.059	0.09
350	1550	150	600	30	300	1.54	0.27	0.34	0.103	0.15
350	1550	150	600	30	300	2.2	0.37	0.49	0.140	0.21
350	1550	150	600	30	300	2.2	0.37	0.27	0.140	0.21

proposed in current study and it is significant difference. which is shown in Table 1. Also the difference between values proposed Priestley (2003) and that obtained from using Eqn. (3) was also significant, because the values proposed by Priestley (2003) not consider the parameter to incorporate the effect of b_w/b_f ratio. To summaries this chapter it may be said that the equations proposed in current study for effective stiffness determination for RC flanged beam sections were verified by comparing with equations developed by other researchers and some of experimental data obtained from various literatures. Some of problems or drawbacks were found in equation proposed by Khuntia and Ghosh (2004) and Priestley (2003) for effective stiffness determination of RC flanged beam sections are discussed in detailed and mentioned. The equations developed in current study estimates the values of effective stiffness of reinforced concrete flanged beam section correctly and hence are valid.

V. Conclusions

The main points of conclusion can be summarized as follows:

1. The present work analyzed the reinforced concrete flanged beam section subjected to lateral loading with taking into account a cracking effect of reinforced concrete structure and proposed the expression to estimate the effective stiffness accurately.
2. Due to cracking, the main dominating factor affecting the nonlinear dependence of structure deformation from lateral load is stiffness reduction of reinforced concrete members.
3. Because of reduction in stiffness (EI) of structure, the lateral deflections of structure may exceed the limit given in seismic code and also the reduction in stiffness (EI) it directly affect on base shear, natural frequency, natural time period of structure, force distribution within members of structure and dynamic response of structure.

4. The stiffness (EI) of reinforced concrete flanged beam sections primarily depends on percentage of steel in section, D_f/D ratio, b_w/b_f ratio and grade of concrete and steel, which has significant influence on effective stiffness flanged beam sections.
5. The expression of the current work is also simplified method of the estimation of structure deflection will allow a quick evaluation of the behaviour of cracked structure under lateral loads.

Acknowledgement

Authors are thankful to the management and principal of Amrutvahini College of Engineering and MIT College of Engineering, for availing the infrastructure and software. They are highly obliged to entire faculty and staff members of Civil Engineering Department.

References

- [1] Ahmed, M., Dad Khan, M.K. and Wamiq, M., Effect of Concrete Cracking on the Lateral Response of RCC Buildings, *Asian Journal of Civil Engineering*, Vol. 9, 2008, pp 25-34.
- [2] Ghosh, S.K. and Khuntia, M., Flexural Stiffness of Reinforced Concrete Columns and Beams: Analytical approach, *ACI Structural Journal*, Vol. 101, 2004, pp 364-374.
- [3] Ghosh, S.K. and Khuntia, M., Flexural Stiffness of Reinforced Concrete Columns and Beams: Experimental Verification, *ACI Structural Journal*, Vol. 101, 2004, pp 351-363.
- [4] Priestley, M.J.N., Myths and Fallacies in Earthquake Engineering, Revisited, The 9th Mallet Milne Lecture, IUSS Press, Pavia, Italy, 2003, pp 1-121.
- [5] Kenneth, J.E. and Marc, O.E., Effective Stiffness of Reinforced Concrete Columns, *ACI Structural Journal*, Vol. 106, 2009, pp 476-483.
- [6] Karar, I.F. and Dundar, C., Effect of Loading Types and Reinforcement Ratio on an Effective Moment of Inertia and Deflection of Reinforced Concrete Beam, *Advances in Engineering Software*, Vol. 40, 2009, pp 836-846.
- [7] IS 1893 (2002), Criteria for Earthquake Resistant Design of Structures, Bureau of Indian Standards, New Delhi, India.
- [8] IS 456 (2000), Indian Standard Plain and Reinforced Concrete – Code of Practice, Bureau of Indian Standards, New Delhi, India.
- [9] IS 13920 (1993), Ductile Detailing of Reinforced Concrete Structures Subjected to Seismic Forces – Code of Practice, Bureau of Indian Standards, New Delhi, India.
- [10] SAP 2000 (V14), Structural Analysis Program - Advanced, Static and Dynamic Finite Element Analysis of Structures, *Computers and structures*, Inc., Berkeley, U.S.A.

Implementation of Low Power and High Speed Multiplier-Accumulator Using SPST Adder and Verilog

H. S. Krishnaprasad Puttam¹, P. Sivadurga Rao² & N. V. G. Prasad³

1, 2,3Department Of E.C.E,Sasi Institute Of Technology & Engineering,Tadepalligudem.India.
1M.Tech Student, 2Assistant Professor & 3Associate Professor,H.O.D.

Abstract: In this paper, we proposed a new architecture of multiplier -and- accumulator (MAC) for high-speed arithmetic and low power. Multiplication occurs frequently in finite impulse response filters, fast Fourier transforms, discrete cosine transforms, convolution, and other important DSP and multimedia kernels. The objective of a good multiplier and accumulator (MAC) is to provide a physically compact, good speed and low power consuming chip. To save significant power consumption of a VLSI design, it is a good direction to reduce its dynamic power that is the major part of total power dissipation. In this paper, we propose a high speed MAC adopting the new SPST implementing approach. This multiplier and accumulator is designed by equipping the Spurious Power Suppression Technique (SPST) on a modified Booth encoder which is controlled by a detection unit using an AND gate. The modified booth encoder will reduce the number of partial products generated by a factor of 2. The SPST adder will avoid the unwanted addition and thus minimize the switching power dissipation.

Keywords: Booth encoder, computer arithmetic, digital signal processing, spurious power suppression technique, low power.

I. INTRODUCTION

With the recent rapid advances in multimedia and communication systems, real-time signal processing's like audio signal processing, video/ image processing, or large-capacity data processing are increasingly being demanded.

The multiplier and multiplier-and-accumulator (MAC) [1] are the essential elements of the digital signal processing such as filtering, convolution, and inner products. Most digital signal processing methods use nonlinear functions such as discrete cosine transform (DCT) [2] or discrete wavelet transform (DWT) [3]. Because they are basically accomplished by repetitive application of multiplication and addition, the speed of the multiplication and addition arithmetic's determines the execution speed and performance of the entire calculation. Because the multiplier requires the longest delay among the basic operational blocks in digital system, the critical path is determined by the multiplier, in general. For high-speed multiplication, the modified radix-4 Booth's algorithm (MBA) [4] is commonly used. However, this cannot completely solve the problem due to the long critical path for multiplication [5], [6].

Power dissipation is recognized as a critical parameter in modern VLSI design field. To satisfy MOORE'S law and to produce consumer electronics goods with more backup and less weight, low power VLSI design is necessary.

Fast multipliers are essential parts of digital signal processing systems. The speed of multiply operation is of great importance in digital signal processing as well as in the general purpose processors today, especially since the media processing took off. In the past multiplication was generally implemented via a sequence of addition, subtraction, and shift operations. Multiplication can be considered as a series of repeated additions. The number to be added is the multiplicand, the number of times that it is added is the multiplier, and the result is the product. Each step of addition generates a partial product. In most computers, the operand usually contains the same number of bits. When the operands are interpreted as integers, the product is generally twice the length of operands in order to preserve the information content. This repeated addition method that is suggested by the arithmetic definition is slow that it is almost always replaced by an algorithm that makes use of positional representation. It is possible to decompose multipliers into two parts. The first part is dedicated to the generation of partial products, and the second one collects and adds them. The basic multiplication principle is twofold i.e., evaluation of partial products and accumulation of the shifted partial products. It is performed by the successive additions of the columns of the shifted partial product matrix. The 'multiplier' is successfully shifted and gates the appropriate bit of the 'multiplicand'. The delayed, gated instance of the multiplicand must all be in the same column of the shifted partial product matrix. They are then added to form the product bit for the particular form. Multiplication is therefore a multi operand operation. To extend the multiplication to both signed and unsigned numbers, a convenient number system would be the representation of numbers in two's complement format.

II. OVER VIEW OF MAC

In this section, basic MAC operation is introduced. A multiplier can be divided into three operational steps. The first is radix-2 Booth encoding in which a partial product is generated from the multiplicand (X) and the multiplier (Y). The second is adder array or partial product compression to add all partial products and convert them into the form of sum and carry. The last is the final addition in which the final multiplication result is produced by adding the sum and the

carry. If the process to accumulate the multiplied results is included, a MAC consists of four steps, as shown in Fig. 1, which shows the operational steps explicitly.

General hardware architecture of this MAC is shown in Fig. 2. It executes the multiplication operation by multiplying the input multiplier (X) and the multiplicand (Y). This is added to the previous multiplication result (Z) as the accumulation step.

The (N)-bit 2's complement binary number can be expressed as

$$X = -2^{N-1} x_{N-1} + \sum_{i=0}^{N-2} x_i 2^i, x_i \in 0, 1 \quad (1)$$

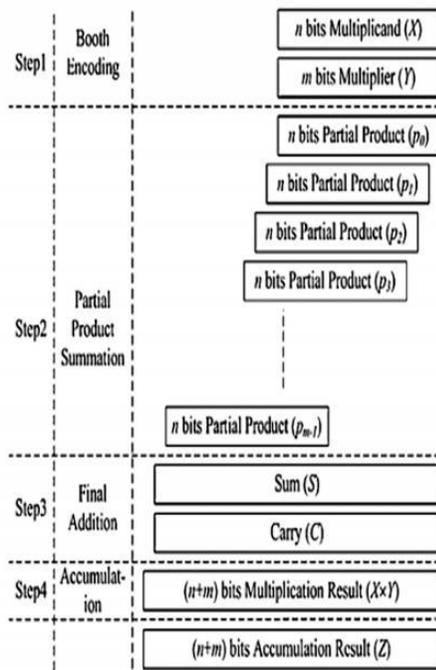


Figure 1: Basic Arithmetic Steps of Multiplication and Accumulation

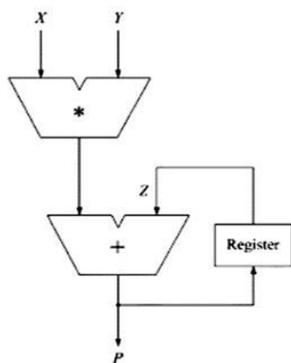


Figure 2. Hardware Architecture of General MAC

If (1) is expressed in base-4 type redundant sign digit form in order to apply the radix-2 Booth's algorithm, it would be [7].

$$X = \sum_{i=0}^{N/2-1} d_i 4_i \quad (2)$$

$$d_i = -2x_{2i+1} + x_{2i} + x_{2i-1} \quad (3)$$

If (2) is used, multiplication can be expressed as

$$X \times Y = \sum_{i=0}^{N/2-1} d_i 2^{2i} Y. \quad (4)$$

If these equations are used, the afore-mentioned multiplication - accumulation results can be expressed as

$$P = X \times Y \times Z = \sum_{i=0}^{N/2-1} d_i 2^{2i} Y + \sum_{J=0}^{2N-1} z_j 2^j. \quad (5)$$

Each of the two terms on the right-hand side of (5) is calculated independently and the final result is produced by adding the two results. The MAC architecture implemented by (5) is called the standard design [6]. If N -bit data are multiplied, the number of the generated partial products is proportional to. In order to add them serially, the execution time is also proportional to N . The architecture of a multiplier, which is the fastest, uses radix-2 Booth encoding that generates partial products.

III. MODIFIED BOOTH ENCODER

In order to achieve high-speed, multiplication algorithms using parallel counters, such as the modified Booth algorithm has been proposed, and some multipliers based on the algorithms have been implemented for practical use. This type of multiplier operates much faster than an array multiplier for longer operands because its computation time is proportional to the logarithm of the word length of operands.

Booth multiplication is a technique that allows for smaller, faster multiplication circuits, by recoding the numbers that are multiplied. It is possible to reduce the number of partial products by half, by using the technique of radix-4 Booth recoding [8]. The basic idea is that, instead of shifting and adding for every column of the multiplier term and multiplying by 1 or 0, we only take every second column, and multiply by ± 1 , ± 2 , or 0, to obtain the same results. The advantage of this method is the halving of the number of partial products.

To Booth recode the multiplier term, we consider the bits in blocks of three, such that each block overlaps the

previous block by one bit. Grouping starts from the LSB, and the first block only uses two bits of the multiplier. Figure 3 shows the grouping of bits from the multiplier term for use in modified booth encoding.

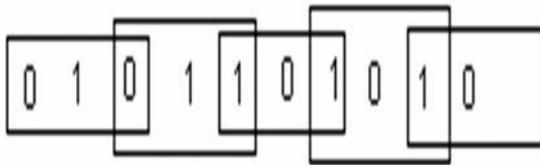


Figure 3. Grouping of Bits from the Multiplier Term

Each block is decoded to generate the correct partial product. The encoding of the multiplier Y , using the modified booth algorithm, generates the following five signed digits, $-2, -1, 0, +1, +2$. Each encoded digit in the multiplier performs a certain operation on the multiplicand, X , as illustrated in Table 1.

Table 1
 Encoded Five Signed Digits

Block	Re-code digit	Operation on X
001	0	0X
001	+1	+1X
010	+1	+1X
011	+2	+2X
100	-2	-2X
101	-1	-1X
110	-1	-1X
111	0	0X

IV. SPURIOUS POWER SUPPRESSION TECHNIQUE

The former SPST has been discussed in [9] and [10]. Figure 4 shows the five cases of a 16-bit addition in which the spurious switching activities occur. The 1st case illustrates a transient state in which the spurious transitions of carry signals occur in the MSP though the final result of the MSP are unchanged. The 2nd and the 3rd cases describe the situations of one negative operand adding another positive operand without and with carry from LSP, respectively. Moreover, the 4th and the 5th cases respectively demonstrate the addition of two negative operands without and with carry-in from LSP. In those cases, the results of the MSP are predictable Therefore the computations in the MSP are useless and can be neglected. The data are separated into the Most Significant Part (MSP) and the Least Significant Part (LSP). To know whether the

MSP affects the computation results or not. We need a detection logic unit to detect the effective ranges of the inputs.

$$A_{MSP} = A [15:8] \quad (6)$$

$$B_{MSP} = B [15:8] \quad (7)$$

$$A_{and} = A [15] \cdot A [14] \dots A [8] \quad (8)$$

$$B_{and} = B [15] \cdot B [14] \dots B [8] \quad (9)$$

$$A_{nor} = A [15] + A [14] + A [13] + \dots + A [8] \quad (10)$$

$$B_{nor} = B [15] + B [14] + B [13] + \dots + B [8] \quad (11)$$

$$Close = \sim ((A_{and} + A_{nor}) \cdot (B_{and} + B_{nor})) \quad (12)$$

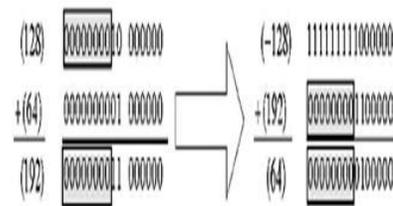
We derive the Karnaugh maps which lead to the Boolean equations (13) and (14) for the Carr_ctrl and the sign signals, respectively. In equation (13) and (14), C_{LSP} denotes the carry propagated from the LSP circuits.

$$Carr_ctrl = (C_{LSP} \oplus A_{and} \oplus B_{and}) (A_{and} + A_{nor}) (B_{and} + B_{nor}) \quad (13)$$

$$Sign = C_{LSP} \cdot \bar{A}_{and} \cdot B_{and} + C_{LSP} \cdot A_{and} \cdot \bar{B}_{and} \quad (14)$$

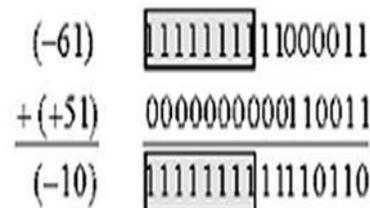
Case 1:

$$(A_0 A_1 \dots A_8) = (1 1 \dots 1), (B_0 B_1 \dots B_8) = (0 0 \dots 0), C_9 = 1$$



Case 2:

$$(A_{15} A_{14} \dots A_8) = (11 \dots 1), (B_{15} B_{14} \dots B_8) = (00 \dots 0), C_7 = 0$$



Case 3:

$$(A_{15} A_{14} \dots A_8) = (11\dots 1), (B_{15} B_{14} \dots B_8) = (00\dots 0), C_7 = 1$$

$$\begin{array}{r} (-196) \quad 1111111100111100 \\ + (204) \quad \boxed{00000000}11001100 \\ \hline (+8) \quad \boxed{00000000}00001000 \end{array}$$

Case 4:

$$(A_{15} A_{14} \dots A_8) = (11\dots 1), (B_{15} B_{14} \dots B_8) = (11\dots 1), C_7 = 0$$

$$\begin{array}{r} (-61) \quad \boxed{11111111}11000011 \\ + (-205) \quad 1111111100110011 \\ \hline (-266) \quad \boxed{11111111}011110110 \\ \quad \quad \quad \underbrace{\quad \quad \quad}_{\text{sign}} \quad \uparrow \text{carr-ctrl} \end{array}$$

Case 5:

$$(A_{15} A_{14} \dots A_8) = (11\dots 1), (B_{15} B_{14} \dots B_8) = (11\dots 1), C_7 = 1$$

$$\begin{array}{r} (-196) \quad 1111111100111100 \\ + (-52) \quad \boxed{11111111}11001100 \\ \hline (-248) \quad \boxed{11111111}00001000 \end{array}$$

Figure 4: Spurious Transition Cases in Multimedia/DSP Processing

V. PROPOSED SPURIOUS POWER SUPPRESSION TECHNIQUE

The SPST uses a detection logic circuit to detect the effective data range of arithmetic units, e.g., adders or multipliers. When a portion of data does not affect the final computing results, the data controlling circuits of the SPST latch this portion to avoid useless data transitions occurring inside the arithmetic units. Besides, there is a data asserting control realized by using registers to further filter out the useless spurious signals of arithmetic unit every time when the latched portion is being turned on. This asserting control brings evident power reduction. Figure 5 shows the design of low power adder/subtractor with SPST.

The adder /subtractor is divided into two parts, the most significant part (MSP) and the least significant part (LSP). The MSP of the original adder/subtractor is modified to include detection logic circuits, data controlling circuits, sign extension circuits, logics for calculating carry in and carry out signals. The most important part of this study is the design of the control signal asserting circuits, denoted as asserting circuits in Figure 5. Although this asserting circuit brings evident power reduction, it may induce additional delay. There are two implementing approaches for the control signal assertion circuits. The first implementing approach of control signal assertion circuit is using registers. This is illustrated in Figure 6. The three output signals of the detection logic are close, Carr_ctrl, sign. The three output signals the detection logic unit are given a certain amount of delay before they assert. The delay \square , used to assert the three output signals, must be set in a range of $\square < \square < \square$, where \square denotes the data transient period and \square denotes the earliest required time of all the inputs. This will filter out the glitch signals as well as to keep the computation results correct. The restriction that $\square < \square$ must be greater than \square to guarantee the registers from latching the wrong values of control usually decreases the overall speed of the applied designs.

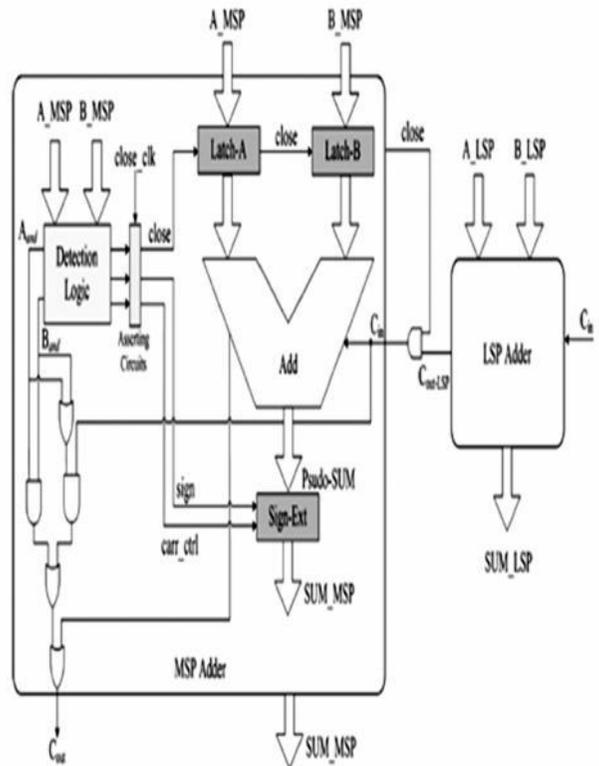


Figure 5. Low Power Adder/Subtractor Adopting the SPST

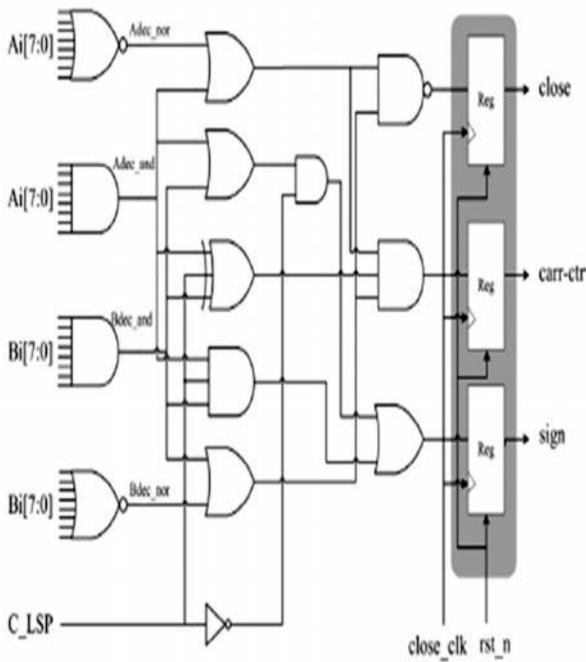


Figure 6. Detection Logic Circuits using Registers

This issue should be noticed in high- end applications which demands both high speed and low power requirements. To solve this problem we adopt the other implementing approach of control signal assertion circuit using AND gate.

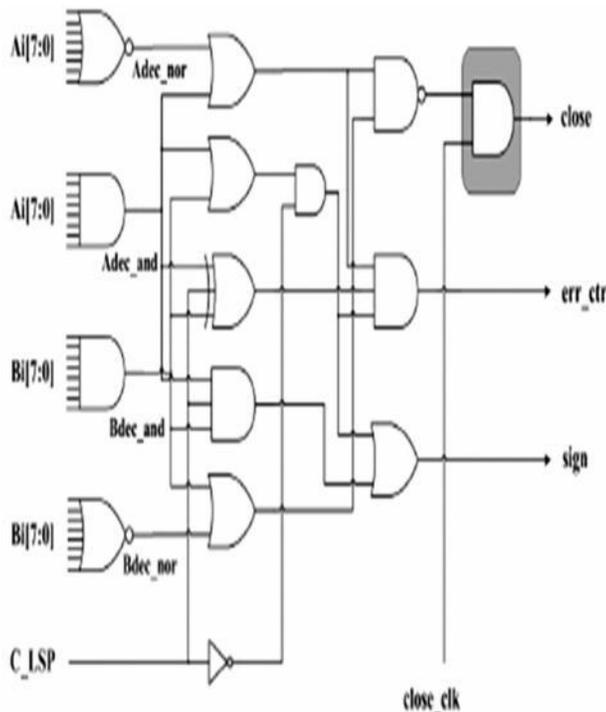


Figure 7. Detection Logic Circuits using AND Gate

VI. PROPOSED LOW POWER HIGH PERFORMANCE MULTIPLIER AND ACCUMULATOR

The proposed high speed low power multiplier is designed by equipping the SPST on a tree multiplier. There are two distinguishing design considerations in designing the proposed multiplier as listed in the following

(A) Applying the SPST on the Modified Booth Encoder

Figure 8 shows a computing example of Booth multiplying two numbers “2AC9” and “006A”. The shadow denotes that the numbers in this part of Booth multiplication are all zero so that this part of the computations can be neglected. Saving those computations can significantly reduce the power consumption caused by the transient signals. According to the analysis of the multiplication shown in figure 8, we propose the SPST-equipped modified-Booth encoder, which is controlled by a detection unit. The detection unit has one of the two operands as its input to decide whether the Booth encoder calculates redundant computations as shown in Figure 9. The latches can, respectively, freeze the inputs of MUX-4 to MUX-7 or only those of MUX-6 to MUX-7 when the PP4 to PP7 or the PP6 to PP7 are zero; to reduce the transition power dissipation. Figure 10. Shows the booth partial product generation circuit. It includes AND/OR/ EX-OR logic.

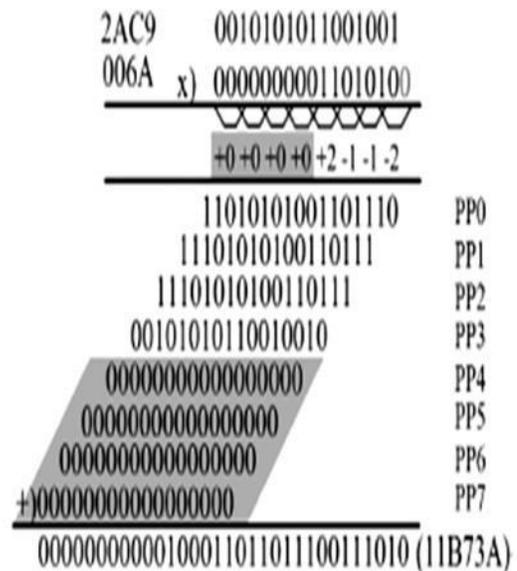


Figure 8: Illustration of Multiplication using Modified Booth Encoding

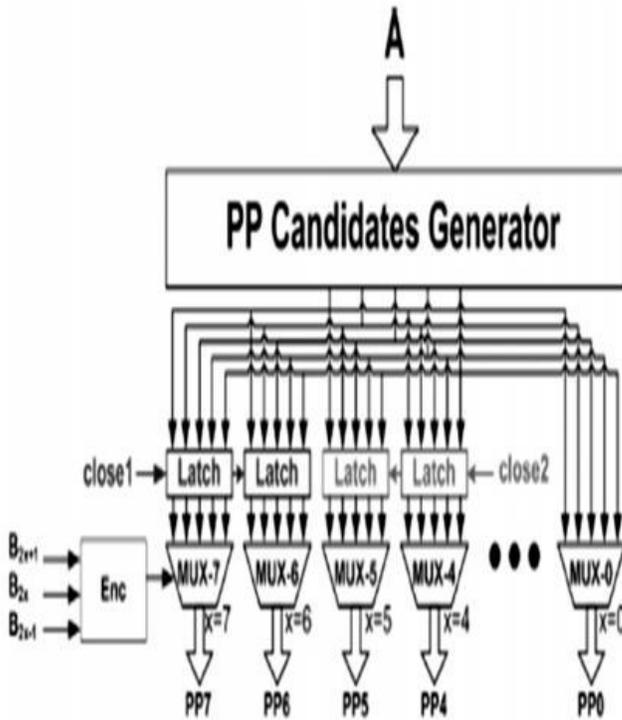


Figure 9: SPST Equipped Modified BoothEncoder

(B) Applying the SPST on the Compression Tree

The proposed SPST -equipped multiplier is illustrated in figure 11. The PP generator generates five candidates of the partial products, i.e., $\{-2A, -A, 0, A, 2A\}$. These are then selected according to the Booth encoding results of the operand B . When the operand besides the Booth encoded one has a small absolute value, there are opportunities to reduce the spurious power dissipated in the compression tree.

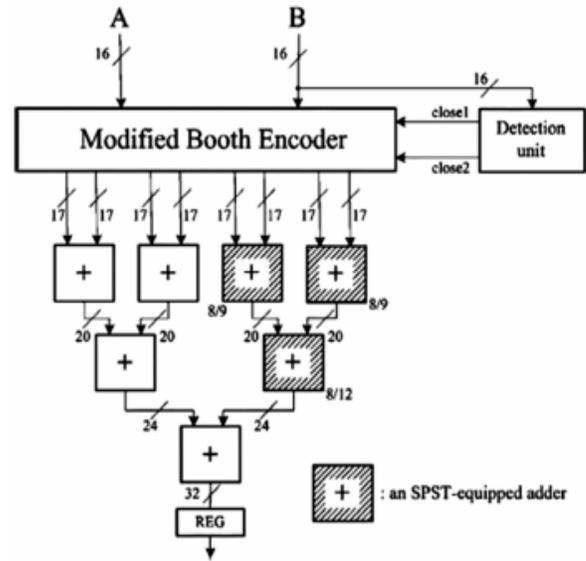


Figure11: Proposed High Performance Low Power Equipped Multiplier

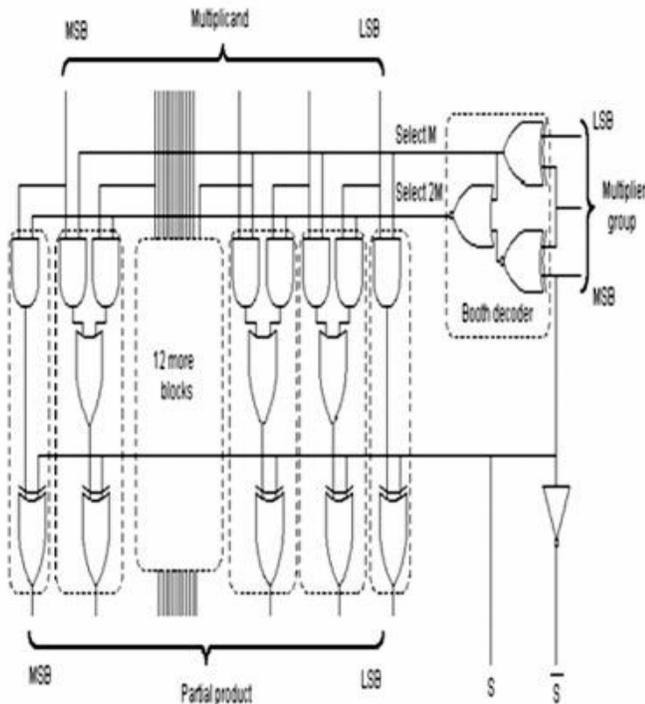


Figure 10: Booth Partial Product Selector Logic

VII. COMPARISON

The Proposed MAC results are very good when compared with the existig methods and the comparision table is given by following table.

Multiplier Type	Array-MAC	MAC-CSA	MAC-SPST
Vendor	Xilinx	Xilinx	Xilinx
Device & Family	Spartan3E XC3s1600e	Spartan3E XC3s1600e	Spartan3E XC3s1600e
Delay	298.974ns	55.916ns	53.805ns
Power Dissipation	123mW	123mW	16mW

VIII. SIMULATION RESULTS

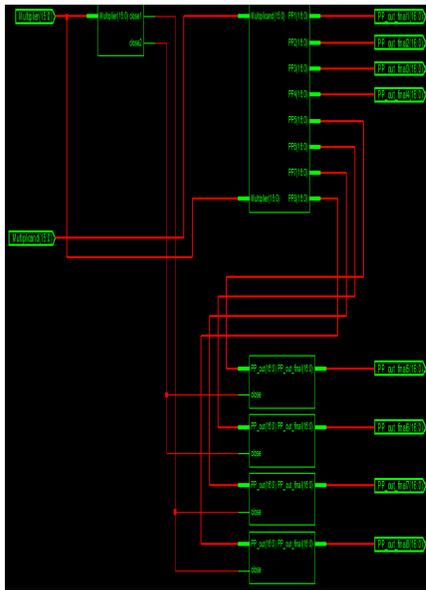


Figure 12: Schematic View of Booth Encoder

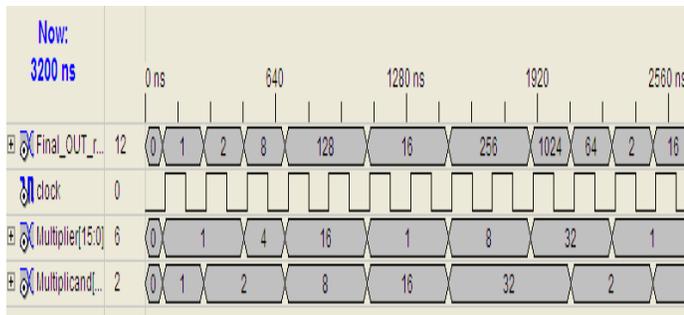


Figure 13: Simulation Waveform of Booth Encoder

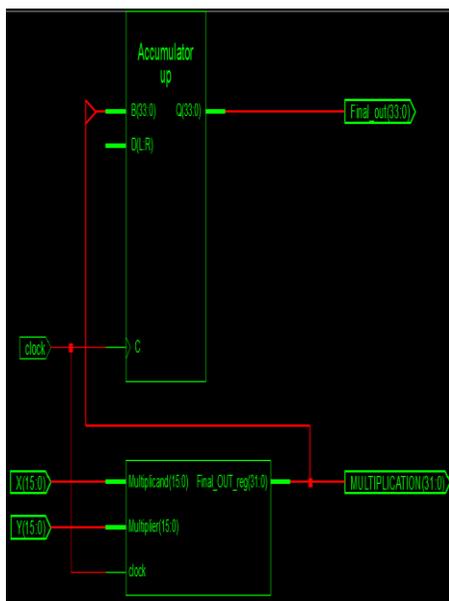


Figure 14: Schematic View of MAC

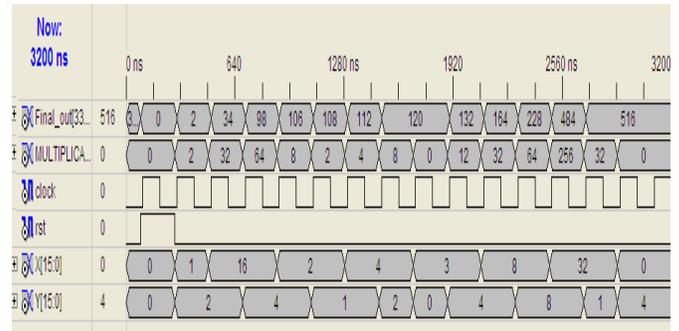


Figure 15: Simulation Waveform of MAC

IX. CONCLUSIONS

In this project, we propose a high speed low-power multiplier and accumulator (MAC) adopting the newSPST implementing approach. This MAC is designed by equipping the Spurious Power Suppression Technique (SPST) on a modified Booth encoder which is controlled by a detection unit using an AND gate. The modifiedbooth encoder will reduce the number of partial products generated by a factor of 2. The SPST adder will avoid the unwanted addition and thus minimize the switching power dissipation. The SPST MAC implementation with AND gates have an extremely high flexibility on adjusting the data asserting time. This facilitates the robustness of SPST can attain 30% speed improvement and 22% power reduction in the modified booth encoder. This design can be verified using Modelsim and Xilinx using verilog.

REFERENCES

- [1] G. Lakshmi Narayanan and B. Venkataramani, "Optimization Techniques for FPGA-Based Wave Pipelined DSP Blocks", *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, **13**, No 7, pp 783-792, July 2005.
- [2] L.Benini, G.D. Micheli, A. Macii, E. Macii, M. Poncino, and R.Scarsi, "Glitching Power Minimization by Selective Gate Freezing", *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, **8**, No. 3, pp. 287-297, June 2000.
- [3] S.Henzler, G. Georgakos, J. Berthold, and D. Schmitt-Landsiedel, "Fastpower-Efficientcircuit-Block Switch-off Scheme", *Electron. Lett.*, **40**, No. 2, pp. 103-104, Jan. 2004.
- [4] H. Lee, "A Power-Aware Scalable Pipelined Booth Multiplier", *In Proc. IEEE Int. SOC Conf.*, 2004, pp. 123-126.
- [5] J. Choi, J. Jeon, and K. Choi, "Power Minimization of Functional units by Partially Guarded Computation", *In Proc. IEEE Int. Symp. Low Power Electron. Des.*, 2000, pp. 131-136.
- [6] Z. Huang and M. D. Ercegovac, "High-Performance

Low-Power Left-Toright Array Multiplier Design”,
IEEE Trans. Comput., **54**, No. 3, pp. 272-283, Mar.
2005.

- [7] K. H. Chen, K. C. Chao, J. I. Guo, J. S. Wang, and Y.S. Chu, “An Efficient Spurious Power Suppression Technique (SPST) and its Applications on MPEG-4 AVC/H.264 Transform Coding Design”, *InProc. IEEE Int. Symp. Low Power Electron. Des.*, 2005, pp. 155-160.
- [8] K. H. Chen, Y. M. Chen, and Y. S. Chu, “A Versatile Multimedia Functional Unit Design using the Spurious Power Suppression Technique”, *in Proc. IEEE Asian Solid-State Circuits Conf.*, 2006, pp. 111-114.
- [9] Zhijun Huang, “High Level Optimization Techniques for Low Power Multiplier Design” University of California, los angels, 2003.

Cutting Speed and Feed Rate Optimization for Minimizing Production Time of Turning Process

S. S. K. Deepak

(Department of Mechanical Engineering, Rungta Engineering College, Raipur, Chhattisgarh, India)

ABSTRACT: Optimum selection of cutting conditions importantly contribute to the increase of productivity by minimization of production time and the associated costs, therefore utmost attention is paid to this problem in this contribution. Time is the most important parameter in any operation and all the manufacturing firms aim at producing a product in minimum time to reach the customer quickly and enhance the customer satisfaction. This can be achieved by using optimization techniques. The success of an optimization technique does not lie in its complexity but the time in which it provides a solution to the manufacturing firms. In this research paper, a geometric programming based approach to optimize the production time of the turning process within in some operating constraints is proposed. It involves mathematical modeling for production time of turning process, which is expressed as a function of the cutting parameters which include the cutting speed and feed rate. Then, the developed mathematical model was optimized with in some operating constraints like the maximum cutting speed, maximum feed rate, power constraint and the surface roughness constraint. The results of the model reveal that the proposed method provides a systematic and efficient technique to obtain the optimal cutting parameters that will minimize the production time of turning process. Thus, it is possible to optimize the production time of turning process effectively by geometric programming. The approach is suitable for fast determination of optimum cutting parameters during machining, where there is not enough time for deep analysis.

Keywords - cutting parameters, geometric programming, model.

Nomenclature:

$$C_{01} = \text{constant} = \left(\frac{\pi dl}{1000} \right)$$

$$C_{02} = \text{constant} = \left(\frac{\pi dl t_c}{1000 Z^n} \right)$$

C_{11} = constant

C_0 = machine cost per unit time (\$/min.)

C_m = machining cost per piece (\$/piece)

C_t = tool cost (\$/cutting edge)

d = diameter of the work piece (mm.)

f = feed rate (mm/revolution)

F = cutting Force (N)

l = length of the work piece (mm.)

n and p are constants.

P_t = production time per piece (min./piece)

R = nose radius of the tool (mm)

R_a = average surface roughness (μm)

t_c = tool changing time (min.)

t_h = tool handling time (min.)

$$t_m = \text{time required to machine a work piece} = \frac{\pi dl}{1000 vf} \text{ (min.)}$$

T = tool life (min.)

v = the cutting speed (m/min.)

Z = constant

η = efficiency of cutting

$\lambda_{01}, \lambda_{02}$ and λ_{11} are lagrange multipliers.

I. Introduction

Turning is a widely used machining process in manufacturing. Therefore, an optimal selection of cutting parameters to satisfy an economic objective within the constraints of turning operation plays a very important role [1]. To determine the optimal cutting parameters, reliable mathematical models have to be formulated to associate the cutting parameters with the cutting performance. However, it is also well known that reliable mathematical models are not easy to obtain [2-5]. In any optimization problem, it is very crucial to identify the prime objective called as the objective function or optimization criterion. In manufacturing processes, the most commonly used objective function is the specific cost [6]. Walvekar and Lambert [7] used geometric programming for the selection of machining variables. The optimum values of both cutting speed and feed rate were found out as a function of depth of cut in multi-pass turning operations [8]. Wu et. al. [9] analyzed the problem of optimum cutting parameters selection by finding out the optimal cutting speed which satisfies the basic manufacturing criterion. Basically, this optimization procedure, whenever carried out, involves partial differentiation for the minimization of unit cost, maximization of production rate or maximization of profit rate. These manufacturing conditions are expressed as a function of cutting speed. Then, the optimum cutting speed is determined by equating the partial differentiation of the expressed function to zero. This is not an ideal approach to the problem of obtaining an economical metal cutting. The other cutting variables, particularly the feed rate also have an important effect on cutting economics. Therefore, it is necessary to optimize the cutting speed and feed rate simultaneously in order to obtain an economical metal cutting operation. The process of the metal cutting depends upon the features of tools, input work materials and machine parameter settings influencing process efficiency and output quality characteristics or responses. A significant improvement in process efficiency may be obtained by process parameters optimization that identifies and determines the regions of critical process control factors leading to desired outputs or responses with acceptable variations ensuring a lower cost of manufacturing [10].

The aim of this research paper is the construction of a mathematical model describing the objective function in terms of the cutting parameters with some operating constraints, then; the mathematical model was optimized by

using geometric programming approach. The developed model and program can be used to determine the optimal cutting parameters to satisfy the objective of obtaining minimum production time of turning process under different operating constraints. The results of the mathematical model are obtained by using suitable software. This research paper proposes a very simple, effective and efficient way of optimizing the production time of the turning process with in some operating constraints such as the maximum cutting speed, maximum feed rate, power requirement, surface roughness. This paper also highlights the merits of geometric programming optimization over other optimization approaches

II. Mathematical Modeling for Optimization

The production time to produce a part by turning operation can be expressed as follows:

$$P_t = \text{Machining Time} + \text{Tool Changing Time} + \text{Set-up Time} \quad (1)$$

$$P_t = t_m + (t_c) \frac{t_m}{T} + t_h \quad (2)$$

The Taylor's tool life (T) used in Eq. (2) is given by:

$$T = \left(\frac{Z}{f^p v} \right)^{\frac{1}{n}} \quad (3)$$

Where, n, p and Z depend on the many factors like tool geometry, tool material, work piece material, etc.

On substituting Eq. (3) in Eq. (2), we get

$$P_t(v, f) = C_{01} v^{-1} f^{-1} + C_{02} v^{\left(\frac{1}{n}\right)-1} f^{\left(\frac{p}{n}\right)-1} + t_h \quad (4)$$

t_h does not depend on cutting speed or feed rate. So, the modified objective function from Eq. (4) can be written as:

$$P_t(v, f) = C_{01} v^{-1} f^{-1} + C_{02} v^{\left(\frac{1}{n}\right)-1} f^{\left(\frac{p}{n}\right)-1} \quad (5)$$

2.1 Machining constraints:

There are many constraints which affect the selection of the cutting parameters. These constraints arise due to various considerations like the maximum cutting speed, maximum feed rate, power limitations, surface finish, surface roughness, etc.

2.1.1 Maximum cutting speed:

The increasing of cutting speed also increases the tool wear, therefore, the cutting speed has to be kept below a certain limit called the maximum cutting speed.

$$v \leq v_{max}. \quad (6)$$

$$C_{11} v \leq 1 \quad (7)$$

$$\text{Where } C_{11} = \frac{1}{v_{max}}. \quad (8)$$

By the method of primal and dual programming of geometric programming, the maximum value of dual function or the minimum value of primal function is given by:

$$v(\lambda) = \left[\frac{C_{01}}{\lambda_{01}} (\lambda_{01} + \lambda_{02}) \right]^{\lambda_{01}} \left[\frac{C_{02}}{\lambda_{02}} (\lambda_{01} + \lambda_{02}) \right]^{\lambda_{02}} \left[\frac{C_{11}}{\lambda_{11}} \right]^{\lambda_{11}} \quad (9)$$

Subject to the following constraints:

$$\lambda_{01} + \lambda_{02} = 1 \quad (10)$$

$$-\lambda_{01} + \left\{ \left(\frac{1}{n} \right) - 1 \right\} \lambda_{02} + \lambda_{11} = 0 \quad (11)$$

$$-\lambda_{01} + \left\{ \left(\frac{p}{n} \right) - 1 \right\} \lambda_{02} = 0 \quad (12)$$

And the non-negativity constraints are:

$$\lambda_{01} \geq 0, \lambda_{02} \geq 0 \text{ and } \lambda_{11} \geq 0 \quad (13)$$

On adding Eq. (10) and Eq. (12), we get

$$\lambda_{02} = n \quad (14)$$

From Eq. (10) and Eq. (14), we get

$$\lambda_{01} = 1 - n \quad (15)$$

From Eq. (11), (14) and (15), we get

$$\lambda_{11} = 1 - p \quad (16)$$

Therefore, the maximum value of dual function and the minimum value of primal function is given by:

$$v(\lambda) = \left(\frac{C_{01}}{1-n} \right)^{1-n} \left(\frac{C_{02}}{n} \right)^n \{C_{11}(1-p)\}^{1-p} \quad (17)$$

$$\text{Now, } \lambda_{11} = \frac{C_{11} v}{v(\lambda)} \quad (18)$$

From Eq. (17) and Eq. (18), we get the optimum values of cutting speed as:

$$v = \frac{\lambda_{11} \left(\frac{C_{01}}{1-n} \right)^{1-n} \left(\frac{C_{02}}{n} \right)^n (C_{11})^{1-p}}{C_{11}} \quad (19)$$

$$\text{And, } \lambda_{01} = \frac{C_{01} v^{-1} f^{-1}}{v(\lambda)} \quad (20)$$

Therefore, from (19) and (20), we get optimum feed rate as:

$$f = \frac{(C_{01} \times C_{11})}{(1-n) \times \lambda_{11} \times \left(\frac{C_{01}}{1-n} \right)^{1-n} \left(\frac{C_{02}}{n} \right)^n (C_{11})^{1-p}} \quad (21)$$

2.1.2 Maximum feed rate:

In rough machining operations, feed rate is taken as a constraint to achieve the maximum production rate.

$$f \leq f_{max}. \quad (22)$$

$$C_{11} f \leq 1 \quad (23)$$

$$\text{Where } C_{11} = \frac{1}{f_{max}}. \quad (24)$$

Following the same procedure as described for the first constraint, we get the following values:

$$\lambda_{01} = 1 - n \quad (25)$$

$$\lambda_{02} = n \quad (26)$$

$$\lambda_{11} = 1 - p \quad (27)$$

$$f = \frac{\lambda_{11} \left(\frac{C_{01}}{1-n} \right)^{1-n} \left(\frac{C_{02}}{n} \right)^n (C_{11})^{1-p}}{C_{11}} \quad (28)$$

$$v = \frac{(C_{01} \times C_{11})}{(1-n) \times \lambda_{11} \times \left(\frac{C_{01}}{1-n} \right)^{1-n} \left(\frac{C_{02}}{n} \right)^n (C_{11})^{1-p}} \quad (29)$$

2.1.3 Power constraint

The maximum power available for the turning operation will be a constraint in the turning operation, which has to be taken in to consideration. The power available for the turning operation is given by:

$$P = \frac{F \times v}{6120 \eta} \leq P_{max}. \quad (30)$$

$$C_{11} v \leq 1 \quad (31)$$

$$\text{Where } C_{11} = \frac{F}{6120 \eta P_{max}}. \quad (32)$$

Following the same procedure as described for the first constraint, we get the following values:

$$\lambda_{01} = 1 - n \quad (33)$$

$$\lambda_{02} = n \quad (34)$$

$$\lambda_{11} = 1 - p \quad (35)$$

$$v = \frac{\lambda_{11} \left(\frac{C_{01}}{1-n}\right)^{1-n} \left(\frac{C_{02}}{n}\right)^n (C_{11})^{1-p}}{C_{11}} \quad (36)$$

$$f = \frac{(C_{01} \times C_{11})}{(1-n) \times \lambda_{11} \times \left(\frac{C_{01}}{1-n}\right)^{1-n} \left(\frac{C_{02}}{n}\right)^n (C_{11})^{1-p}} \quad (37)$$

2.1.4 Surface roughness

Surface roughness can be used as a constraint in finishing operations. Therefore, it becomes a very important factor in determining finish cutting conditions. Surface roughness can be expressed in terms of feed as follows:

$$R_a = \frac{f^2}{32R} \quad (38)$$

$$R_a C_{11} \leq \frac{f^2}{32R} \quad (39)$$

Following the same procedure as described for the first constraint, we get the following values:

$$\lambda_{01} = 1 - n \quad (40)$$

$$\lambda_{02} = n \quad (41)$$

$$\lambda_{11} = \frac{(1-p)}{2} \quad (42)$$

$$v = \frac{\left[\frac{(1-p)}{2}\right] \left[\frac{C_{01}}{1-n}\right]^{1-n} \left[\frac{C_{02}}{n}\right]^n [C_{11}]^{1-p}}{C_{11}} \quad (43)$$

$$f = \frac{C_{01} \times C_{11}}{(1-n) \left\{ \frac{(1-p)}{2} \right\} \left[\frac{C_{01}}{1-n} \right]^{1-n} \left[\frac{C_{02}}{n} \right]^n [C_{11}]^{1-p}} \quad (44)$$

III. Results and Discussion

3.1 Figures:

The figures obtained from the implementation of the mathematical model are as follows:

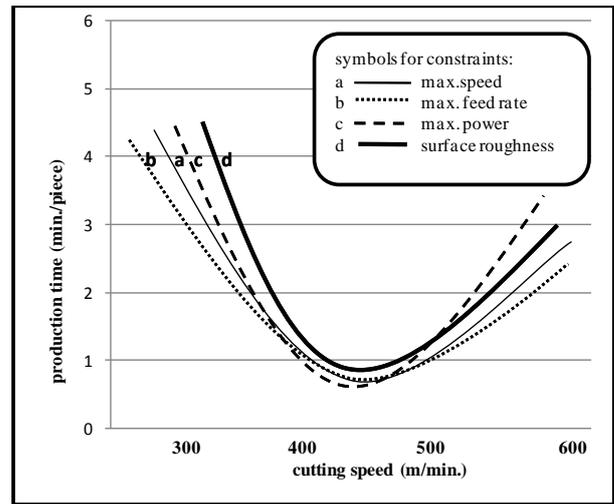


Fig 1: Production time versus cutting speed curve for different constraints.

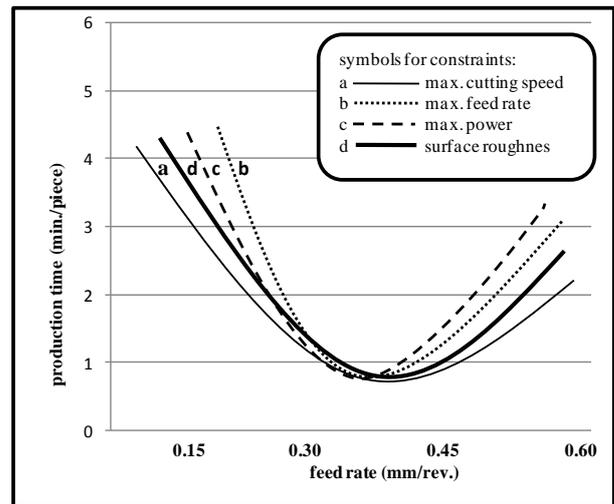


Fig 2: Production time versus feed rate curve for different constraints.

3.2 Analysis of results:

The curves obtained between production time and cutting speed reveal that a smaller value of cutting speed results in a high production time. It is due to the fact that a smaller cutting speed increases the production time of parts. Also, it will decrease the profit rate due to the production of a lesser number of parts. However, if the cutting speed is too high, it will also lead to a high production time due to excessive tool wear and increased machine downtime. The optimum cutting speed is somewhere between “too slow” and “too fast” which will yield the minimum production time.

The curves between the production time and the feed rate indicate that a small feed rate will result in high production time. A smaller feed rate means the number of revolutions should be increased. The more the number of revolutions, the more will be the production time. Even a very high feed rate is not advisable as it will increase the tool wear and surface roughness resulting in increased machining time and machine downtime resulting in high production time. So, the optimum feed rate is somewhere between “too small” and “too high” which will result in the minimum production time.

IV. Conclusion

In this research paper, the cutting speed and feed rate were modeled for the minimum production time of a turning operation. The maximum cutting speed, the maximum feed rate, maximum power available and the surface roughness was taken as constraints. The results of the model show that the proposed method provides a systematic and efficient method to obtain the minimum production time for turning. This approach helps in quick analysis of the optimal region which will yield a small production time rather than focusing too much on a particular point of optimization. It saves a lot of time and can be easily implemented by manufacturing firms. The developed model will provide with the optimal values of the cutting speed and feed rate that will satisfy the objective of production time minimization within the given operating constraints. The coefficients n , p and Z of the extended Taylor's tool life equation are not described in depth for all cutting tool and work piece combinations. Obtaining these coefficients experimentally requires lot of time, resources and then, the analysis of the obtained values increases the complexity of the process.

It can be concluded from this study that the obtained model can be used effectively to determine the optimum values of cutting speed and feed rate that will result in minimum production time. The developed model saves a considerable time in finding the optimum values of the cutting parameters. It has been shown that the method of geometric programming can be applied successfully to optimize the production cost of turning process.

REFERENCES

- [1] Gilbert W.W., Economics of machining in theory and practice, *American society for metal, Cleveland*, 1950, 465-485.
- [2] A.M. Abuelnaga, M.A. El-Dardiry, Optimization methods for metal cutting, *International Journal of Machine Tool Design Research*, 1984, 24 (1), 11-18.
- [3] B. White, A. Houshyar, Quality and optimum parameter selection in metal cutting, *Compu. Ind.*, 1992, 20, 87- 98.
- [4] C. Zhou, R.A. Wysk, An integrated system for selecting optimum cutting speeds and tool replacement times, *International Journal of Machine Tools Manufacturing*, 1992, 32(5), 695-707.
- [5] M.S. Chua, M. Rahman, Y.S. Wong, H.T. Loh, Determination of optimal cutting conditions using design of experiments and optimization techniques, *International Journal of Machine Tools Manufacturing*, 1993, 33 (2), 297-305.
- [6] Liang M., Mgwatu, M., Zuo, M., Integration of cutting parameter selection and tool adjustment decisions for multi-pass turning, *International Journal of Advanced Manufacturing Technology*, 2001, 17, 861-869.
- [7] Tan, F.P., and Creese, R.C., A generalized multi-pass machining model for machining parameter selection in turning, *International Journal of Production Research*, 1995, 33 (5), 1467-1487.
- [8] Walvekar, A.G., and Lambert, B.K., An application of geometric programming to machining variable selection, *International Journal of Production Research*, 1970, 8, 241-245.
- [9] Wu, S.M., and Emer, D.S., Maximum profit as the Criterion in Determination of Optimum Cutting Conditions, *Journal of Engineering for Industry*, 1966, 1, 435-442.
- [10] Montgomery, D.C., 1990, *A textbook on statistical quality control*, 2nd edition.

Analysis of Design Level Defects Based On PetriNet Model

P. Rajesh¹, P. Rajarajeswari², Dr. D. Vasumathi³

* (M.Tech, Computer Science and Engineering, Madanapalle Institute of Technology & Sciences, Madanapalle, A.P,India)

** (M.Tech[Ph.d], Assistant Professor, Department of CSE, Madanapalle Institute of Technology & Sciences, Madanapalle, A.P, India)

*** (M.Tech, Ph.d, Associate Professor, Department of CSE, JNTU Hyderabad, Hyderabad, A.P,India)

Abstract: Software Designs must be evaluated in Software Development Process so as to avoid bugs and unsatisfactory performances. Concerning performance and real-time properties modeling, OMG specified the UML Profile for Schedulability, Performance and Time Specification (SPT). UML is the standard OO modeling Language our system, but UML is too static to model the performance. It is not able to capture dynamic nature of system. So here we are using PetriNets to capture Dynamic nature of the System for UML use Case Diagrams and collaboration diagrams. First we draw the UML Use Case diagrams and collaboration diagrams with SPT (i.e. performance information) and then convert them into Executable Petri Net models. Finally we consider a case study for our proposed algorithm.

Keywords: Software Performance Engineering, UML, Petri Nets, Performance Evaluation.

I. INTRODUCTION

Performance is an important but often overlooked aspect of the software design. Indeed, the consideration on performance issues is in many cases left until late in the software development process (SDP), when problems have already manifested themselves at system test or within the deployed system. The identification of possible bugs or unsatisfactory performance in the design phase allows to contain the costs, also permitting to compare different alternatives. This kind of approach implements so called software performance engineering (SPE)[2], which is a systematic, quantitative technique to construct software systems that meet performance objectives.

Using the UML for modeling and the OMG UML Profile for Schedulability, Performance and Time Specification (SPT)[9] to specify performance requirements into a UML model therefore mapped into a performance model (Petri Nets). In this perspective, the main contribution of this paper is the implementation of a software development process which takes into account performance specifications and requirements: the software performance engineering development process (SPEDP). SPEDP includes and synthesizes both the aim of modeling and developing generic software architecture, and the aim of investigating the performance of the overall (hardware/software) elaboration system. It can be applied both in early phases, as a software performance engineering technique, and in test phases, as a common software performance evaluation technique. SPEDP fixes steps, rules and guidelines to follow in order to achieve the desired results in the software development satisfying the performance requirements.

UML is too static to model dynamic behavior of the systems. So to overcome this we use Petri Nets as to develop Executable models.

Bernardi et al. have proposed the automatic translation of state charts and Sequence Diagrams into Generalized Stochastic Petri Nets, as well as a composition of the resulting net models suitable for reaching a given analysis goal [3]. Elkoutbi et al. have transformed a simple use case structure to colored Petri nets [4] and Kamandi et al. have transformed use case to Object Stochastic Activity Network (OSAN) [5]. Different approaches are used for the transformation of sequence diagrams to Petri nets. In the approach proposed by Bernardi et al., all structures of the sequence diagram have been transformed to Generalized Stochastic Petri Nets [3]. Ourdani et al. transformed the simplest structures in the sequence diagram to colored Petri nets [6]. The difference between the two transformations is that in Bernardi et al.'s approach [3] the transformation is based on mapping messages as well as conveying them, while in Ourdani et al.'s approach [6] the transformation is based on message sender and receiver component.

Although so many researchers have used this Petri Nets as performance Domain, none of the Researchers utilized the collaboration diagrams and Use Case diagrams for performance evaluation [3, 5] of software design description based on Petri Nets either in this paper we propose an algorithm for transforming annotated collaboration diagrams to Petri Nets by referring [1].

While Use Case diagrams model the functions of system components Collaboration diagrams model the interaction between the system components (i.e. the messages exchanged between the components). These diagrams enriched by performance input parameters. Then we must transform these input parameters to tokens of places or guards for arcs and transitions in the Petri Net Model. After that target model must be evaluated. The designers will decide whether and how software architecture should be refined from analysis of the results of the analysis of the results of the evaluation steps.

The rest of the paper is organized as follows. In section 2 an algorithm for transmission of the collaboration diagrams and Use Case enriched by performance parameters will be proposed. Section 3 presents a case study which illustrates the proposed algorithm. Ultimately the section 4 concludes the paper.

II. THE TRANSFORMATION OF UML DIAGRAMS TO PETRI NETS

In the following, the collaboration diagrams and role of them concerning performance will be explained. Also we use UML Profile for Schedulability, Performance and Time (SPT) to annotate the additional information to these diagrams. Then the detail of the transformation algorithm will be explained.

A. Role of the collaboration Diagram concerning performance

Collaboration diagrams are useful design tools because they provide a dynamic view of the system behavior, which can be difficult to extract from static diagrams or specifications. To each message in the diagram a condition can be attached, representing the possibility that the message could be dispatched. Even multiple messages can leave a single point each one labeled by a condition. From the performance point of view it can be considered that routing rates are attached to the messages. We use PApob tag value to give such information. A set of messages can be dispatched multiple times if they are enclosed and marked as an iteration. This construction also has its implications from the performance point of view [7]. We use the PApob tag value of the PAstep stereotype to annotate the probability of execution an alternative behavior when the sequence diagram presents either alt, break or option fragment operator. Also iteration will be represented by the tag value PArep [7]. The PADemand tag value specifies the duration of the activities as random variables exponentially distributed, which are the ones supported by the extension of Petri nets formalism.

Also PADelay tag value specifies the delay of the messages exchanged among components allocated in different physical nodes [7].

B. The Transformation of Collaboration diagrams to Petri Nets

-Petri nets representation of asynchronous and delayed messages: We use PADemand tag value to give such information. In this state, the client component is displayed in the form of place-transition-place. But the server component is shown as place-transition-place-transition- place. The second transition is a timed transition with an assigned firing rate. Figure 2.a shows this type of messages.

Petri nets representation of synchronous and delayed messages: We use PADemand tag value to give such information. In this state the client component is shown in the form of place-transition-place-transition-place and their connection can happens through two shared places. But the server component is displayed as place-transition-place-transition-place-transition-place where central transition is of the timed transition. Figure 2.b shows this type of message.

Figure 2.c depicts the message exchanged between two annotated components c1 and c2 with PADelay tag value and the resulting Petri net, where t1 represents the sending action performed by component c1, t2 models the message transmission delay and t3 represents the reception of the message by component c2. The value associated to the tag PADelay defines the firing rate of the timed transition, t2.

Figure 3 shows the two types of sequence diagram constructors (alternative and loop) and their mapping onto Petri nets. The translation of these constructors requires the use of additional Petri net sub-nets. Figure 3.a2 shows the Petri net sub-net modeling, the alternative choice between ev1 and ev2. The additional sub-nets are enclosed in the dotted rectangle. Figure 3.b2 shows the Petri net modeling as an optional choice. Consider that the choices in these two figures are probabilistic. The weights of the conflicting

transitions t1 and t2 are derived from the tag value PApob attached to the constraint condition. Finally, Figure 3.c2 and 3.c3 demonstrate the Petri net sub-net modeling, a while-do loop and repeat-until, respectively. The sub-nets repeat and while models the iteration of message ev1.

C. Role of the Use case Diagram Concerning Performance

Use Case Diagram describes the software system at a Very high level of abstraction by identifying its functionalities. Thus this type of diagram gives information on the type of traffics incoming in the system. The use case diagram should represent a Performance Context, since it specifies one or more scenarios that are used to explore various dynamic situations involving a specific set of resources. Then, it is stereotyped as PAcontext. Each use case used with performance evaluation purposes could represent a step. Then, they are stereotyped as PAstep. The performance annotations for the use case diagram are the assignment of a probability to every edge that links a type of actor to a use case, i.e. the probability of the actor to execute the use case [7, 8]. We use PApob tag value for this annotation. Also each use case of interest should be detailed by means of a Collaboration diagram.

D. The transformation of Use case Diagrams to Petri Nets

In this paper, in order to transform a use case diagram to a Petri net, the idea presented by Elkoutbi et al. [4] is applied with some adjustment. The annotated use case diagram can be transformed to the Petri net through the following steps:

The transformation of each use case to a Petri net:

In this transformation, each use case is transformed into a Petri Net model. We use one place for each actor and one dark place for each use case. The input for Place is a transition with a guard. In the next stage, the dark places are replaced with the obtained Petri net from the Collaboration diagram. Figure 1 displays an actor and two use cases that are annotated with SPT profiles. Then these diagrams are transformed to equivalent Petri net. The selection condition of each use case is assigned to t1 and t2 transitions.

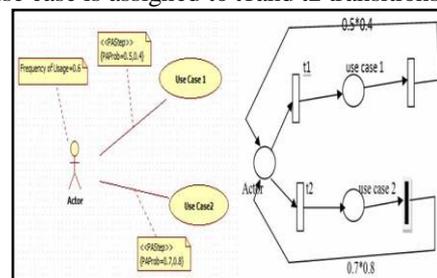


Figure1: A Use Case Diagram and a Petri Net for it.

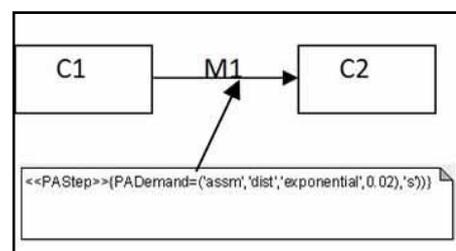


Figure2.1: Component diagram for asynchronous message

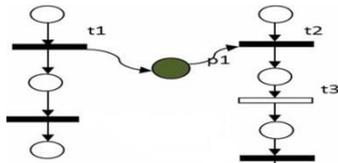


Figure 2.1a: Respective Petri Net for 2.1 annotated PADemand

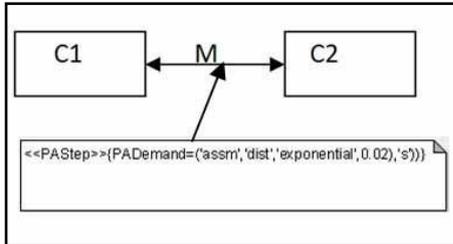


Figure 2.2 : Component diagram of Synchronous Message

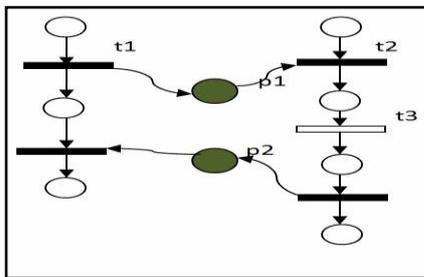


Figure 2.2a: Respective Petri Net for 2.2 annotated PADemand

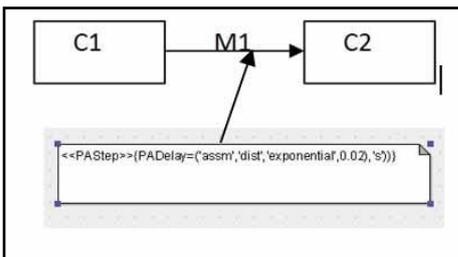


Figure 2.3 : Component Diagram annotated PADelay

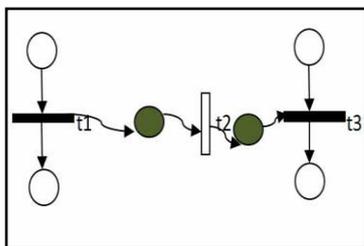
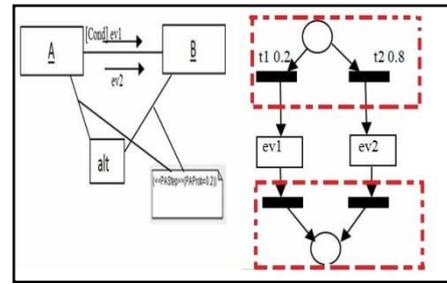


Figure 2.3a: Respective Petri Net for 2.1 annotated PADelay

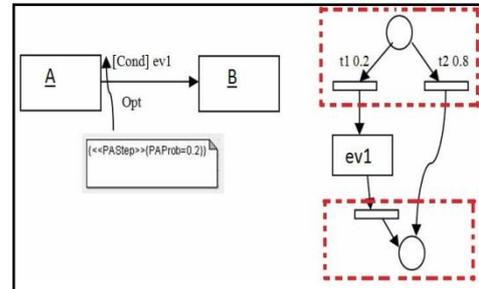
III. CASE STUDY

To represent the usage of our proposed algorithm, in this section we consider a single Automated Teller Machine (ATM) as an example. The use case diagram of an ATM system is shown in Figure 4.a. One of the sequence diagrams of ATM system corresponding to the use case "Identify" is shown in Figure 4.b. This diagram is used when the PIN entered by a customer is valid; the identification will be successfully done.



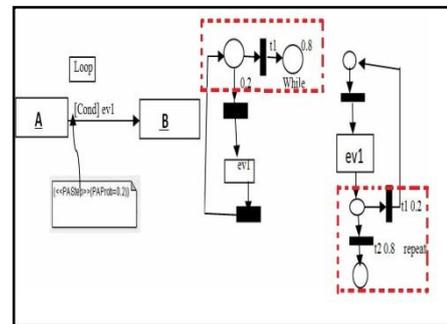
a1

a2



b1

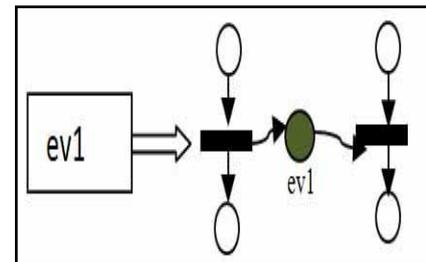
b2



c1

c2

c3



d1

Figure 3: a1, a2, b1, b2, c1, c2, c3, d1 Petri Net Model for Component diagram structures.

According to the proposed algorithm, the equivalent Petri net of these diagrams are shown in Figure 5. ATM, bank, account and customer in Collaboration diagram which are the components in the Petri net model have been presented as separate columns. Each column presents a separate component. In this way, we distinguish the internal arcs which show transformation from one state of a component to another state, and the external arcs which show message exchange between two components. In use case diagram, customer place selects one of these use cases: balance, withdrawal or deposit. In each use case, the identification of the customer is mandatory.

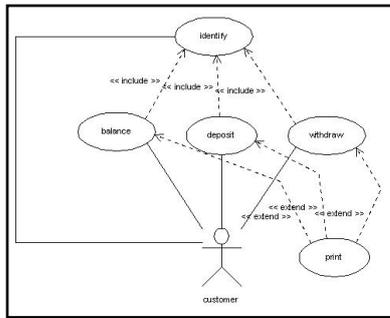


Figure 4(a) : Use Case Diagram for ATM

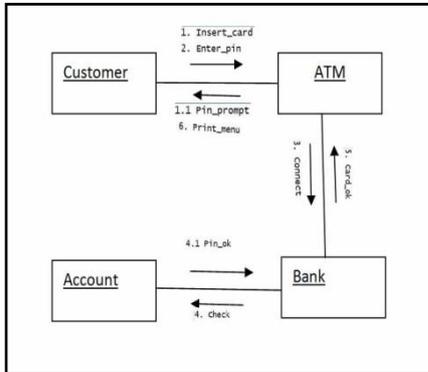


Figure 4(b) : Component Diagram of ATM

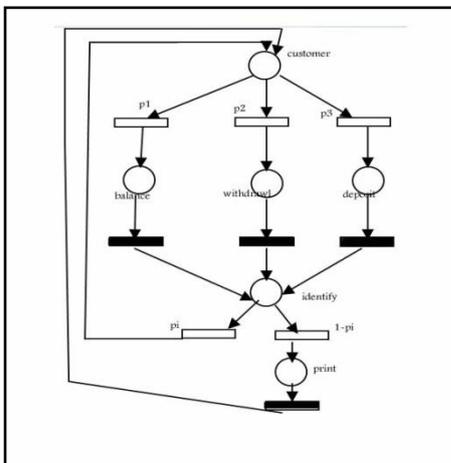


Figure 5a: Petri Net of above Use case diagram(4.a)

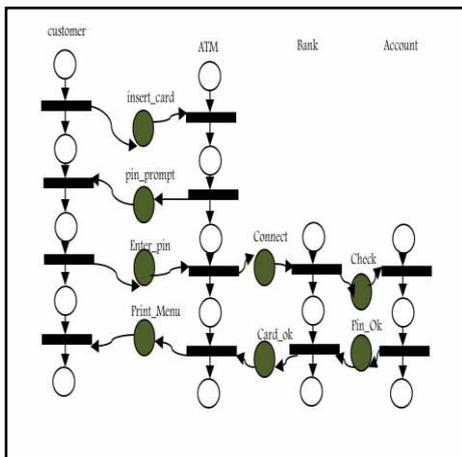


Figure 5(b): Petri Net for above component Diagram

IV. CONCLUSION

In this paper, we transformed annotated use case and collaboration diagrams with performance parameter to Petri net notations. In our further researches, we will consider the transformation of other annotated software architecture description diagrams with performance parameter to an executable model. Moreover, we can consider the annotation of additional information of other non-functional requirements to the software architecture description diagrams, as well. So, the resulting executable model can be used for evaluating those non-functional requirements.

REFERENCES

- [1] Sima Emadi Engineering Department, Maybod Branch, Islamic Azad University, Yazd, Iran and Fereidoon Shams, Computer Engineering Department, Shahid Beheshti University, Tehran- Iran Mapping annotated use case and sequence Diagrams to a Petri Net Notation for Performance Evaluation. (reference).
- [2] C.U. Smith and L.G. Williams, Performance Solutions: A Practical Guide to Creating Responsive, Scalable Software. Addison Wesley Longman Publishing Co., Inc., 2002.
- [3] Bernardi S, Donatelli S, and Merseguer J. 2002, From UML Sequence Diagrams and Statecharts to Analysable Petri Net Model", In Proceedings of the 3rd International Workshop on Software and Performance, ACM, 35-45.
- [4] Elkoutbi M, and Rodulf K. 1998, Modelling Interactive Systems With Hierarchical Coloured Petri Nets, In Proceeding of the Advanced Simulation Technologies Conference, Boston, MA, 432-437.
- [5] Kamandi A, Abdollahi Azgomi M, Movaghar A. 2006, Transformation of UML Models into Analyzable OSAN Models, Electronic Notes in Theoretical Computer Science 159, (Elsevier, 2006), 3-22.
- [6] Ourdani A, Esteban P, Paludetto M and Pascal J. C. A. 2006, Meta Modeling Approach for Sequence Diagram to Petri Nets Transformation Within the Requirements Validation Process, In 20th annual European Simulation and Modelling Conference, LAAS, Toulouse, France, ESM'2006 conference.
- [7] Bernardi S, Merseguer J. 2007, Performance Evaluation of UML Design with Stochastic Well-Formed Nets, The Journal of Systems and Software 80, science direct, 1843-1865.
- [8] Merseguer J, and Campos J. 2003, Exploring Roles for the UML Diagrams in Software Performance Engineering, In proceedings of the 2003 International Conference on Software Engineering Research and Practice (SERP'03), Las Vegas, Nevada, USA: CSREA Press, 43-47.
- [9] Object Management Group, "UML Profile for Schedulability, Performance and Time Specification," OMG, Jan. 2005.

Voice and Speech Recognition for Tamil Words and Numerals

V. S. Dharun, M. Karnan

*Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India.
Prof and Head, Department of Computer Science and Engineering, Tamilnadu College of Engineering, Coimbatore
Tamilnadu, India.*

ABSTRACT: Voice Recognition is often confused with Voice and speech Recognition, which is the translation of spoken words (voice and speech) into text. But the identification of "who" is speaking is not the same activity as the recognition of what words are being spoken.

The main objective of this research is to develop a system for voice recognition in Tamil Word and Numeral using Mel-Frequency Cepstral Coefficients (MFCC) and Dynamic Time Warping (DTW).

Tamil language has 247 letters, but most of them are derived from the 12 vowels and 18 consonants. The other 216 letters are made by combining the sounds of a vowel and a consonant. Each letter is having unique sound.

Voice recognition is the system by which sounds, words or phrases spoken by humans are converted into electrical signals and these signals are transformed into coding patterns to which meaning has been assigned.

To extract valuable information from the voice and speech signal, make decisions on the process, and obtain results, the data needs to be manipulated and analyzed. This research work presents the feasibility of MFCC to extract features and DTW to compare the Tamil words and numerals test patterns. The extraction and matching Process is implemented right after the Pre Processing or filtering signal is performed.

I. INTRODUCTION

For the past forty years, voice and speech recognition research has been characterized by the steady accumulation of small incremental improvements. There has also been a trend to change focus towards more difficult tasks due both to progress in voice and speech recognition performance and to the availability of faster computers.

This research attempts to take advantage of the fact that in many applications there is a large quantity of speech data available, up to millions of hours. It is too expensive to have humans transcribe such large quantities of speech, so the research focus is on developing new methods of machine learning that can effectively utilize large quantities of unlabeled data. Another area of research is better understanding of human capabilities and to use this understanding to improve machine recognition performance.

Because of their limitations and high cost, voice recognition systems have traditionally been used only in a few specialized situations. For example, such systems are useful in instances when the user is unable to use a keyboard to enter data because his or her hands are occupied or disabled. Instead of typing commands, the user can simply speak into a headset. Increasingly, however, as the cost decreases and performance improves, voice and speech recognition systems are entering the mainstream and are being used as an alternative to keyboards

LITERATURE SURVEY

Designing a machine that mimics human behavior, particularly the capability of speaking naturally and responding properly to spoken language, has intrigued engineers and scientists for centuries. The research in automatic voice and speech recognition by machine has attracted a great deal of attention over the past five decades.

- In the late 1960^s, Atal and Itakura independently formulated the fundamental concepts of Linear Predictive Coding (LPC), which greatly simplified the estimation of the vocal tract response from voice and speech waveforms.
- By the mid 1970^s, the basic ideas of applying fundamental pattern recognition system to voice and speech recognition, based on LPC methods, were proposed by Itakura, Rabiner and Levinson and others.
- Another system that was introduced in the late 1980^s was the idea of Artificial Neural Networks (ANN).

Digital processing of voice and speech signal and voice recognition technique is very important for fast and accurate automatic voice recognition system. The voice is a signal of Infinite information. A direct analysis and synthesizing the complex voice signal is tough due to too much information contained in the signal. Therefore the digital signal processes such as Feature Extraction and Feature Matching are introduced to represent the voice signal.

Linear Predictive Coding (LPC), Hidden Markov Model (HMM), Artificial Neural Network (ANN) and etc are evaluated with a view to identify a straight forward and effective method for voice signal. The Voice is a signal of infinite information. Nowadays it is being used for health care, telephony military and people with disabilities therefore the digital signal processes such as Feature Extraction and Feature Matching are the latest issues for study of voice signal.

The most common approaches to voice recognition can be divided into two phases: "template matching" and "feature analysis". Template matching is the simplest technique and has the highest accuracy when used properly, but it also suffers from the most limitations. As with any approach to voice recognition [1-7].

II. METHODOLOGY

The types of voice and speech differences that the speaker-independent method can deal with, but which pattern matching would fail to handle, include accents, and varying speed of delivery, pitch, volume, and inflection. Speaker-independent voice and speech recognition has proven to be very difficult, with some of the greatest hurdles being the variety of accents and inflections used by speakers of different nationalities. Recognition accuracy for speaker-independent systems is somewhat less than for speaker-dependent systems, usually between ninety and ninety five percent. Voice Recognition based on the speaker can be classified into two types namely: Speaker-dependent and Speaker-independent.

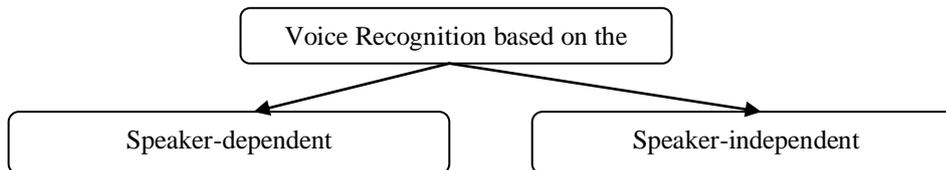


Fig1: Voice Recognition based on the speaker

Voice Recognition based on the speaker words can be classified into three types namely: Discrete, Connected and Continuous speech

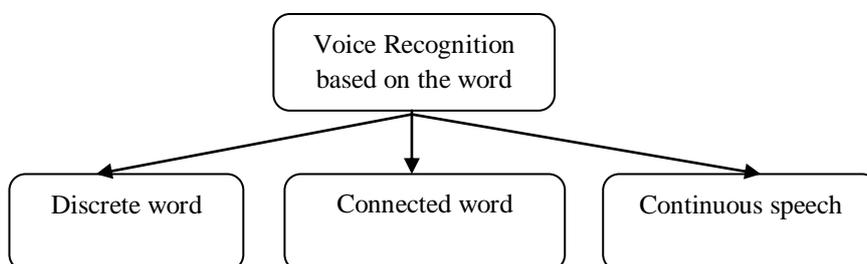


Fig 2: Voice Recognition based on the speaker words

III. DISCRETE WORD

Another way to differentiate between voice recognition systems is by determining if they can handle only discrete words, connected words, or continuous voice and speech. Most voice recognition systems are discrete word systems, and these are easiest to implement. For this type of system, the speaker must pause between words. This is fine for situations where the user is required to give only one word responses or commands, but is very unnatural for multiple word inputs.

IV. CONNECTED WORD

In a connected word voice recognition system, the user is allowed to speak in multiple word phrases, but he or she must still be careful to articulate each word and not slur the end of one word into the beginning of the next word. Totally natural, continuous voice and speech includes a great deal of "co articulation", where adjacent words run together without pauses or any other apparent division between words. A voice and speech recognition system that handles continuous voice and speech is the most difficult to implement.

Speech system is there are two types namely Discrete and Continuous

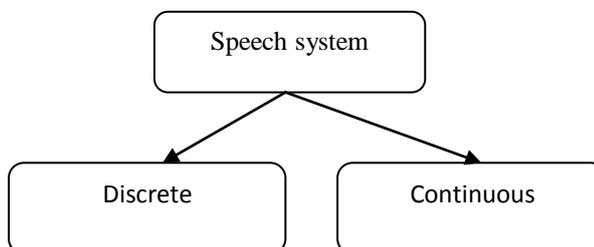


Fig3: Speech system

The template matching method of voice recognition is based on the general principles of digital electronics and basic computer programming. To fully understand the challenges of efficient speaker-independent voice recognition, the fields of phonetics, linguistics, and digital signal processing should also be explored.

The most powerful systems can recognize thousands of words. However, they generally require an extended training session during which the computer system becomes accustomed to a particular voice and accent. Such systems are said to be speaker dependent. Many systems also require that the speaker speak slowly and distinctly and separate each word with a short pause. These systems are called discrete voice and speech systems. Recently, great strides have been made in continuous voice and speech systems voice recognition systems that allow you to speak naturally.

- The voice recognition is available through feature analysis and this technique usually leads to speaker-independent voice recognition.
- Instead of trying to find an exact or near-exact match between the actual voice input and a previously stored voice template.
- This method first processes the voice input using Fourier transforms or Linear Predictive Coding (LPC), then attempts to find characteristic similarities between the expected inputs and the actual digitized voice input. These similarities will be present for a wide range of speakers, and so the system need not be trained by each new user.

OVERVIEW

- The first phase is for the user to speak a word or phrase into a microphone.
- The electrical signal from the microphone is digitized by an Analog-to-Digital (A/D) converter, and is stored in memory.
- To determine the "meaning" of this voice input, the computer attempts to match the input with a digitized voice sample, or template that has a known meaning. This technique is a close analogy to the traditional command inputs from a keyboard.
- The program contains the input template, and attempts to match this template with the actual input using a simple conditional statement.

NUMERAL RECOGNITION SYSTEM

There are several kinds of parametric representation of the acoustic signals. Among them the Mel-Frequency cepstral Coefficient (MFCC) is most widely used. We have developed the recognition system using MFCC and DTW

TRAINING

Since each person's voice is different, the program cannot possibly contain a template for each potential user, so the program must first be "trained" with a new user's voice input before that user's voice can be recognized by the program. During a training session, the program displays a printed word or phrase, and the user speaks that word or phrase several times into a microphone.

The program computes a statistical average of the multiple samples of the same word and stores the averaged sample as a template in a program data structure. With this approach to voice recognition, the program has a "vocabulary" that is limited to the words or phrases used in the training session, and its user base is also limited to those users who have trained the program. This type of system is known as speaker dependent. It can have vocabularies on the order of a few hundred words and short phrases, and recognition accuracy can be about ninety five percent.

FEATURE EXTRACTION AND MATCHING

Feature extraction techniques are emerging such as pith-synchronous signal analysis, phase spectrum based acoustic features, etc. The optimization of the existing features may draw attention

Feature matching in this phase, MFCC coefficients of both the voice and speech signals are compared using the concept of Dynamic Time Warping. This technique is for measuring similarity between two time series which may vary in time or speed. This technique also used to find the optimal alignment between two times series if one time series may be "warped" non-linearly by stretching or shrinking it along its time axis.

One of the earliest approaches to isolated word voice and speech recognition was to store a prototypical version of each word (called a template) in the vocabulary and compare incoming voice and speech with each word, taking the closest match.

Comparing the template with incoming voice and speech might be achieved via a pair wise comparison of the feature vectors in each. The total distance between the sequences would be the sum or the mean of the individual distances between feature vectors. The problem with this approach is that if constant window spacing is used, the lengths of the input and stored sequences are unlikely to be the same.

The Dynamic Time Warping technique achieves this goal; it finds an optimal match between two sequences of feature vectors which allows for stretched and compressed sections of the sequence. A single recording of each word is used as the basis for the stored template in a DTW based recognizer. This approach will not be very robust since it takes no account of the variability of different utterances and does not ensure that the template is representative of the class as a whole.

A voice of Tamil word and numeral analysis is done after taking an input through microphone from a user. The design of the system, involves manipulation of the input audio signal. At different levels, different operations are performed on the input signal such as Pre-emphasis, Framing, Windowing, Mel Cepstrum analysis and Recognition (Matching) of the spoken word.

The research work is to build a Tamil word and numeral recognition tool for Tamil language. This is an isolated word voice and speech recognition tool.

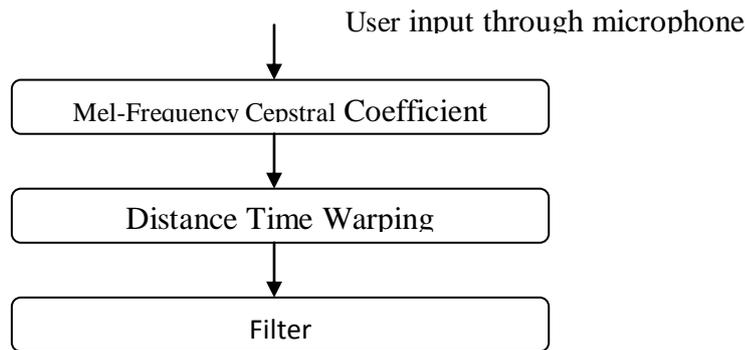


Fig4: Feature Extraction

This work has discussed two phases used for voice recognition system which are important in improving its performance.

- First phase provides the information, to extract MFCC coefficients from the voice signal
- Second phase endow with the technique to compare or match them with the already fed user's voice features using DTW (dynamic time warping technique).

In voice processing, the Mel-Frequency Cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency. Mel-Frequency Cepstral Coefficients (MFCCs) are coefficients that collectively make up an MFC. They are derived from a type of cepstral representation of the audio clip (a nonlinear "spectrum-of-a-spectrum"). The difference between the cepstrum and The Mel-Frequency Cepstrum is that in the MFC, the frequency bands are equally spaced on the Mel scale, which approximates the human auditory system's response more closely than the linearly-spaced frequency bands used in the normal cepstrum.

MFCCs are derived as follows:

- Take the Fourier Transform of (a windowed excerpt of) a signal.
- Map the powers of the spectrum obtained above onto the Mel scale, using triangular overlapping windows.
- Take the logs of the powers at each of the Mel frequencies.
- Take the discrete cosine transform of the list of Mel log powers, as if it were a signal.
- The MFCCs are the amplitudes of the resulting spectrum.

MFCCs are used as features in speech recognition systems, such as the systems which can automatically recognize numbers spoken into a telephone.

A cepstrum is the result of taking the Fourier transform (FT) of the logarithm of the estimated spectrum of a signal. There is a complex cepstrum, a real cepstrum, a power cepstrum, and phase cepstrum. The power cepstrum in particular finds applications in the analysis of human speech. The name "cepstrum" was derived by reversing the first four letters of "spectrum". Operations on cepstra are labeled quefrequency analysis, cepstral analysis. The power cepstrum of a signal is defined as the squared magnitude of the Fourier transform of the logarithm of the squared magnitude of the Fourier transform of a signal.

$$\text{Power cepstrum of signal} = \mathcal{F} \{ \mathbf{Log}(\mathcal{F}\{f(t)\})^2 \}^2$$

The cepstrum is a representation used in homomorphic signal processing, to convert signals (such as a source and filter) combined by convolution into sums of their cepstra, for linear separation.

The result is called the Mel-Frequency Cepstrum or MFC (its coefficients are called Mel-Frequency Cepstral Coefficients, or MFCCs). It is used for voice identification, pitch detection and much more. The independent variable of a cepstral graph is called the quefrequency. The quefrequency is a measure of time, though not in the sense of a signal in the time domain.

DYNAMIC TIME WARPING (DTW)-BASED SPEECH RECOGNITION

Dynamic time warping is an approach that was historically used for speech recognition but has now largely been displaced by the more successful HMM-based approach. Dynamic time warping is an algorithm for measuring similarity between two sequences that may vary in time or speed.

A well-known application has been automatic speech recognition, to cope with different speaking speeds. In general, it is a method that allows a computer to find an optimal match between two given sequences (e.g., time series) with certain restrictions. That is, the sequences are "warped" non-linearly to match each other. This sequence alignment method is often used in the context of hidden Markov models.

Dynamic time warping (DTW) is an algorithm for measuring similarity between two sequences which may vary in time or speed. For instance, similarities in walking patterns would be detected, even if in one video the person was walking slowly and if in another he or she were walking more quickly, or even if there were accelerations and decelerations during the course of one observation. DTW has been applied to video, audio, and graphics indeed, any data which can be turned into a linear representation can be analyzed with DTW. A well known application has been automatic speech recognition, to cope with different speaking speeds.

In general, DTW is a method that allows a computer to find an optimal match between two given sequences (e.g. time series) with certain restrictions. The sequences are "warped" non-linearly in the time dimension to determine a measure of their similarity independent of certain non-linear variations in the time dimension. This sequence alignment method is often used in t This example illustrates the implementation of dynamic time warping when the two sequences are strings of discrete symbols. $d(x, y)$ is a distance between symbols, i.e. $d(x, y) = |x - y|$.

These both techniques have been worked out for same voice and speech signals as well as for different voice and speech signals and it have been found that if both voice and speech signals are same the cost will be 0 and if voice and speech signal are of different voices then cost will definitely have some value which shows the mismatching of the signals.

While voice and speech recognition is the process of converting voice and speech to digital data, voice recognition is aimed toward identifying the person who is speaking. Voice recognition works by analyzing the features of voice and speech that differ between individuals. Everyone has a unique pattern of voice and speech stemming from their anatomy (the size and shape of the mouth and throat) and behavioral patterns.

The applications of voice recognition are markedly different from those of voice and speech recognition. Most commonly, voice recognition system is used to verify a speaker's identity or determine an unknown speaker's identity. Speaker verification and speaker identification are both common types. Speaker verification is the process of using a person's voice to verify that they are who they say they are. Essentially, a person's voice is used like a fingerprint. [9-13]

V. CONCLUSION

The Voice and speech is the most prominent and natural form of communication between humans. There are various spoken the excitation signal is spectrally shaped by a vocal tract Equivalent filter. The outcome of this process is the sequence of exciting signal called voice and speech.

The digitized samples are then processed using MFCC to produce Tamil word and numeral features. After that, the coefficient of Tamil numeral features can go through DTW to select the pattern that matches the database and input frame in order to minimize the resulting error between them. The popularly used cepstrum based methods to compare the pattern to find their similarity are the MFCC and DTW. The MFCC and DTW features can be implemented using MATLAB.

Hamming window is used as window shape by considering the next block in feature extraction processing chain and integrates all the closest frequency lines. Fast Fourier Transform to convert each frame of N samples from time domain into frequency domain FFT is being used. The Fourier Transform is used to convert the convolution of the glottal pulse and the vocal tract impulse response in the time domain.

Discrete Cosine Transform (DCT) this is the process to convert the log Mel spectrum into time domain using DCT. The result of the conversion is called Mel Frequency Cepstrum Coefficient. The set of coefficient is called acoustic vectors. Therefore, each input utterance is transformed into a sequence of acoustic vector. Delta energy and delta spectrum the voice signal and the frames changes, such as the slope of a formant at its transitions. Therefore, there is a need to add features related to the change in cepstral features over time. 13 delta or velocity features (12 cepstral features plus energy), and 39 features a double delta or acceleration feature are added. The energy in a frame for a signal x in a window from time sample t1 to time sample t2, is represented.

Feature matching technique is based on Dynamic Programming and DTW. This technique is used for measuring similarity between two time series which may vary in time or speed. The theoretical and experimental comparisons of the methods are finding an obviously superior feature combination technique.

Experiments and results:

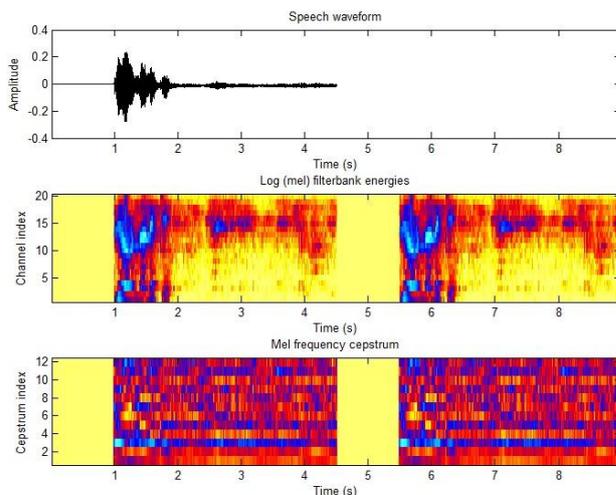


Fig : Speech and MFCC Waveforms for Tamil Numeral "ONDRU"

In the above figure first the wave form of Tamil Numeral "ONDRU" is given which gives the variation of amplitude of speech signal in accordance with time. The second and third plot shows the spectrum of Log (mel) filter bank energies Mel Frequency Cepstrum for Tamil Numeral "ONDRU"

The mel frequency can be approximated by the following equation:

$$mel(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right)$$

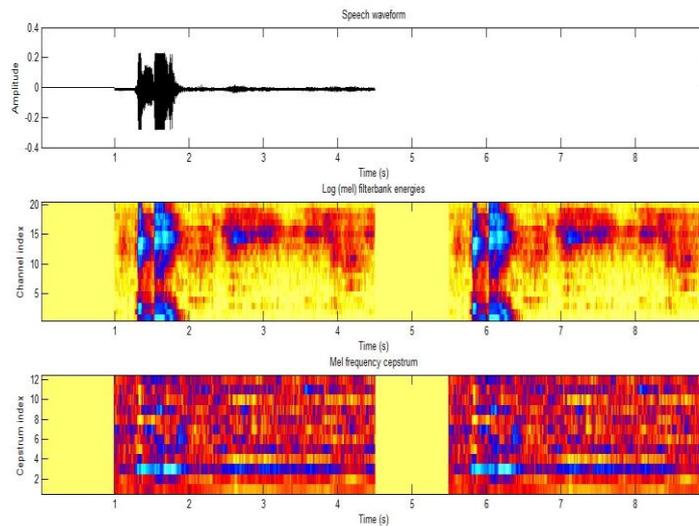
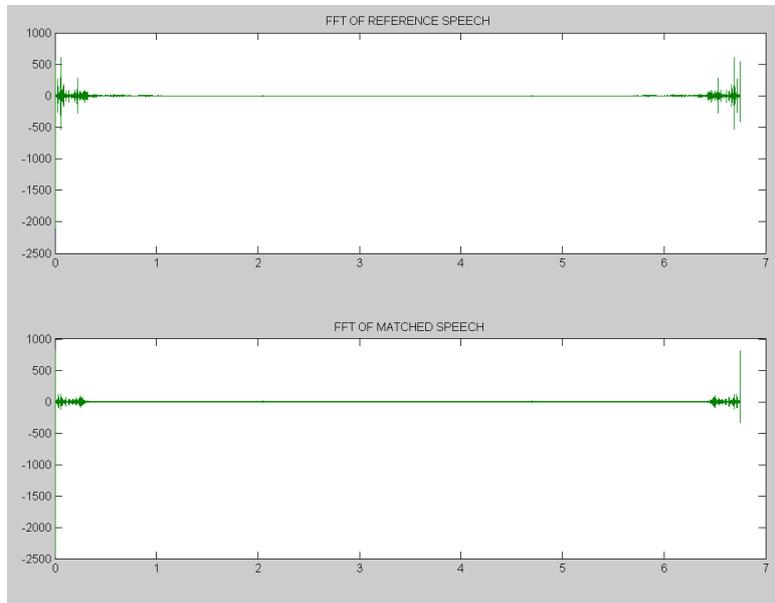


Fig: Speech and MFCC Waveforms for Tamil Numeral “IRANDU”

In the above figure first the wave form of Tamil Numeral “IRANDU” is given which gives the variation of amplitude of speech signal in accordance with time. The second and third plot shows the spectrum of Log (mel) filter bank energies and Mel Frequency Cepstrum for Tamil Numeral “IRANDU”

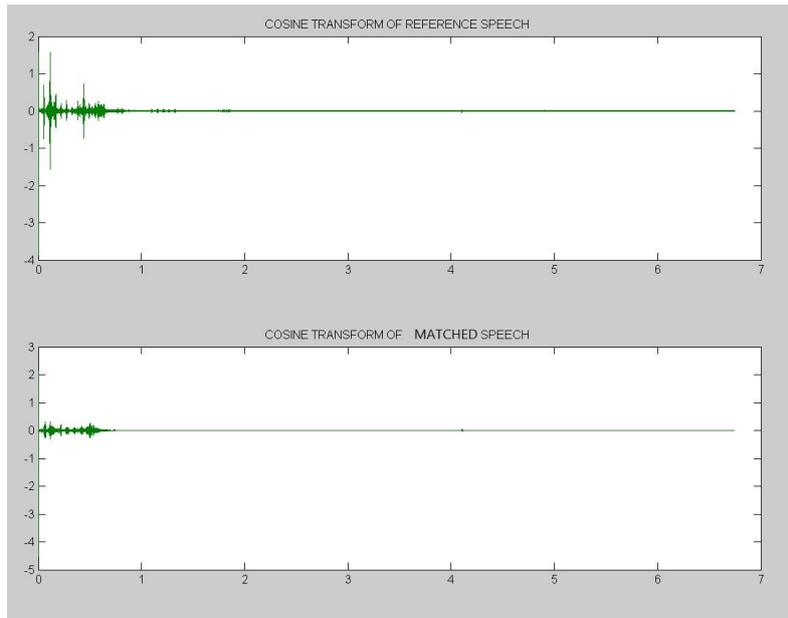
The mel frequency can be approximated by the following equation:

$$mel(f) = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right)$$



The above figure shows the FFT of the Original speech and matched speech. Here the FFT is calculated for spectral estimation of the speech signal. FFT is calculated using the formula

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{2\pi i k n}{N}}, k=0, \dots, N-1$$



The above plot shows the discrete cosine transform of the original speech and the matched speech. The process of discrete cosine transform is carried out before calculating MFCC. DCT can be calculated using the formula

$$c[n] = \sum_{i=1}^M \log(Y(i)) \cdot \cos\left(\frac{\pi n}{M} \left(i - \frac{1}{2}\right)\right)$$

Distance coordinates for “IRANDU” matching speech

0	0	16	4777	9195	16028
16	16	0	5329	7378	14379
4777	4777	5329	0	484	486
9195	6986	7378	484	0	441
16028	13819	12170	486	441	0

The above distance coordinates give the distance between the spoken speech and the reference if the distance coordinate is zero it represents that the speech matches

Output window for “IRANDU” in matching speech



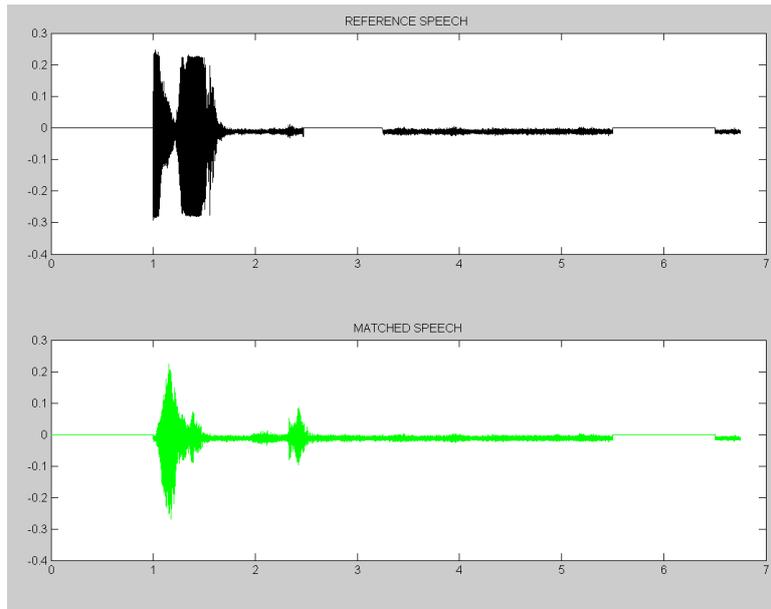
Distance coordinates for “IRANDU” matching speech

1	3	12	4912	9520	16585
18	17	3	5332	7513	14704
4779	4778	5332	3	487	489
9197	6988	7379	487	3	444
16030	13821	12172	489	444	3

Output window for “IRANDU” in matching speech

If the distance coordinate is a non-zero value which means that Spoken speech does not match with the template





the above plot shows the matching speech spoken two different users where the first one is the reference speech and the second one is the word spoken by the speaker. Both the speech are matched by the concept of Dynamic time warping (DTW).

SPEECH BY USER	REFERENCE SPEECH	DISTANCE
ONDRU	PUJJUM	94.968
	ONDRU	0
	IRANDU	116.780
	NANGU	56.743
	AYNTHU	76.873

COMPARISON OF SPEECH BY USER AND REFERENCE SPEECH

SPEECH BY USER	REFERENCE SPEECH	DISTANCE
AMMA	AMMA	0
	APPA	32.981
	AKKA	16.790
	ANNA	36.351
	THAMBHI	154.873

The above table compares the speech spoken by the user and the reference speech templates using DTW and obtains the distance between the speech signals. The table shows that if there is a matching speech the distance is zero else a non zero value turns up as distance value.

REFERENCES

- [1] Adams, Russ, Sourcebook of Automatic Identification and Data Collection, Van Nostrand Reinhold, New York, 1990.
- [2] Cater, John P., Electronically Hearing: Computer Voice and speech Recognition, Howard W. Sams & Co., Indianapolis, IN, 1984.
- [3] Fourcin, A., G. Harland, W. Barry, and V. Hazan, editors, Voice and speech Input and Output Assessment, Ellis Horwood Limited, Chichester, UK, 1989.
- [4] Yannakoudakis, E. J., and P. J. Hutton, Voice and speech Synpaper and Recognition Systems, Ellis Horwood Limited, Chichester, UK, 1987.
- [5] Christopher Hale, CamQuynh Nguyen, "Voice Command Recognition Using Fuzzy Logic" ,Motorola, Austin, Texas 78735, pp 608-613, ISBN no: 0-7803-2636-9.
- [6] Hubert Wassner and Gerard Chollet, "New Time Frequency Derived Cepstral Coefficients For Automatic Voice and speech Recognition", 8thEuropean Signal Processing Conference (Eusipco'96).
- [7] Marco Grimaldi and Fred Cummins, "Speaker Identification Using Instantaneous Frequencies" ,IEEE Transactions On Audio, Voice and speech, And Language Processing, VOL. 16, NO. 6, pp 1097-1111, ISBN: 1558-7916, August 2008.

- [8] Mahdi Shaneh addnd Azizollah Taheri, "Voice Command Recognition System Based on MFCC and VQ Techniques" ,World Academy of Science, Engineering and System 57 2009, pp 534-538.
- [9] Muda Lindasalwa, Begam Mumtaj and Elamvazuthi I., " Voice Recognition Techniques using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", Journal Of Computing, Volume 2, Issue 3, pp 138-143, ISSN 2151-9617, March 2010.
- [10] Norhaslinda Kamaruddin and Abdul Wahab, "Voice and speech Emotion Verification System (Sevs) Based On MFCC For Real TimeApplications", School of Computer Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798.
- [11] Paulraj M P1, Sazali Bin Yaacob1, Ahamad Nazri2 and Sathees Kumar1,"Classification of Vowel Sounds Using MFCC and FeedForward Neural Network" ,5th International Colloquium on Signal Processing & Its Applications (CSPA), pp 60 -63, ISBN: 978-1-4244-4152-5, March 2009.
- [12] Rozeha A. Rashid, Nur Hija Mahalin, Mohd Adib Sarijari and Ahmad Aizuddin Abdul Aziz, "Security System Using Biometric System: Design and Implementation of Voice Recognition System (VRS)",Proceedings of the International Conference on Computer and Communication Engineering, pp 898-902, ISBN :978-1-4244-1692-9, May 2008.
- [13] Suzuki H., Zen H., Nunkuku Y., Miyajima C., Tokuda K., and Kitumuru I.,"Voice and speech Recognition Using Voice-Characteristic dependent Acoustic Models" ,ICASSP 2003,pp 740-743,ISBN: 0-7803-7663-3103,2003.

Determination of Unbalance in Rotating Machine Using Vibration Signature Analysis

B. Kiran Kumar¹, G. Diwakar², Dr. M. R. S. Satynarayana³

¹M. Tech Student, Mech Dept, P.V.P.S.I.T, KANURU, Vijayawada-7, Andhra Pradesh, INDIA

²Assoc. Professor, Mech Dept, P.V.P.S.I.T, KANURU, Vijayawada-7, Andhra Pradesh, INDIA

³Vice principal, Gitam University, Visakhapatnam, Andhra Pradesh, INDIA

ABSTRACT: Vibrations are found almost everywhere in rotating machines. Rotating machinery vibrates due to unbalances, misalignments and imperfect bearings. Vibrational analysis of rotating machinery is able to identify a large number of system ills. Shaft bow, shaft unbalance and coupling misalignments make up the major portion of the observed vibrational frequency spectra of rotating machinery. These vibrational spectra can be used to determine the type of rotating system abnormality. Unbalance is the most cause of machine vibration, an unbalanced rotor always cause more vibration and generates excessive force in the bearing area and reduces the life of the machine. In this paper, experimental studies were performed on a rotor to predict the unbalance in rotor. The vibration velocities were measured at five different speeds using FFT (Fast Fourier Transform) at initial condition. Based on vibration readings spectrum analysis and phase analysis was carried out to determine the cause of high vibrations. By observing the spectrum unbalance was identified. Then Rotor was balanced and found that vibrations were reduced.. The experimental frequency spectra were obtained for both balanced and unbalanced condition under different unbalanced forces at different speed conditions. This paper aims at the implementation of condition based maintenance on rotating machine, by adopting Vibration spectrum analysis which is a predictive maintenance technology. It eliminates unnecessary opening of equipment with considerable savings in personnel resources.

Keyword: Vibration Signature Analysis, Vibration Spectrum Analysis, Unbalance, FFT

Nomenclature:

MNDE: Motor Non Drive End;

H: Horizontal;

v: Velocity

MDE: Motor Drive End;

V: Vertical

PBE: Pillow Block End;

A: Axial

I. INTRODUCTION

Rotor unbalance is the most common reason in machine vibrations. Most of the rotating machinery problem can be solved by using the rotor balancing and misalignment. A very small amount of unbalance may cause severe problem in high speed rotating machines. Overhung rotors are used in many engineering applications like pump, fans, propellers and turbo machinery. The vibration signature of the overhung rotor is totally different from the center hung rotors. The vibration caused by unbalance may destroy critical parts of the machine, such as bearings, seals, gears and couplings. In practice, rotors can never be perfectly balanced because of manufacturing errors such as porosity in casting, non-uniform density of material, manufacturing tolerances and gain or loss of material during operation [1] As a result of mass unbalance, a centrifugal force is generated and must be reacted against by bearing and support structures. A number of analytical methods have been applied to unbalance response such as the transfer method [2]. Further, the unbalance part of the rotor rotates at the same speed as the rotor and therefore the force caused by the unbalance is synchronous [3]. However all the above investigations resulted in fu numerical solutions of the unbalance responses of coupled two-shaft rotor-bearing system. On the other hand, Rao [4] suggested analytical closed-form expressions for the major and minor axis radii of the unbalance response orbit for one-shaft rotor-bearing system. Rao *et al.* [5] and Shiau *et al.*[6]. Vibration signatures are widely used as a useful tool for studying progressive machine mechanical malfunctions, and also form the baseline signature for further comparative monitoring to detect mechanical faults [7].In this paper a general method is presented for obtaining the unbalance response orbit based on the experimental, where the shafts rotate at different speeds. Unbalance system of an overhung rotors are considered for unbalance study. Experiments were conducted for a single mass, at five different speeds and corresponding results are plotted. The rotor unbalance can be detected by spectral and phase analysis.

Description of the Experimental Setup: The Experimental apparatus is shown in photograph of Figure 2 and Figure 3. It consists of a 0.5 hp A.C. Induction motor 1440 rpm speed, a fixed type flange coupling and a single disk rotor. The rotor shaft is supported by single identical ball bearing (pillow block) and has a length of 1000 mm with a bearing span of 750 mm. The diameter of the rotor shaft is 16 mm. A disk of 220 mm in diameter and 6 mm in thickness is mounted on the rotor shaft at bearing end. The rotor shaft is driven by 0.5 hp A.C. motor. The speed of the motor is controlled by using VFD (Variable Frequency Drive) which is mainly used for A.C motors, to increase or decrease the speeds of the motor in the range of 500 to1440 rpm. The instrument used in experiment includes FFT which measures the vibration in terms of velocity at MNDE, MDE & PBE housing and gives the corresponding values.

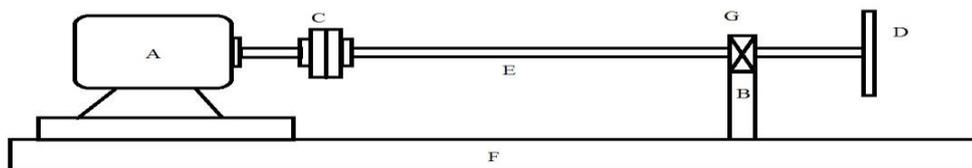


Fig. 1: Line diagram of experimental setup.

A-A.C Induction Motor, B-Bearing Support, C-Coupling, D-Disk, E-Rotor Shaft, F-Base, G-Pillow Block



Fig. 2: Photo graph of VFD



Fig. 3: Photo graph of experimental setup

Experimental Procedure: Experimental facility as shown in Figure: 2 and 3 is used for unbalance test. First the setup is run for few minutes to settle down all minor vibrations. Before creating the unbalanced, the shaft is checked for any misalignment and unbalance. After this an unbalance has been created by placing a mass of 96 gram in the overhung rotor at a radius of 98 mm. FFT is the vibration analyzer is used to acquire the vibration signals in terms of velocity. Vibration signals are measured at five different speeds 600, 800, 1000, 1250 and 1440 rpm with the unbalanced rotor system at drive end (DE), non drive end (NDE) and Pillow block end (PBE) stored in the vibration analyzer. And same masses were added at exact opposite direction i.e in balanced state and vibration signals were taken. Following are the vibration readings observed.

Vibration readings at unbalanced condition:

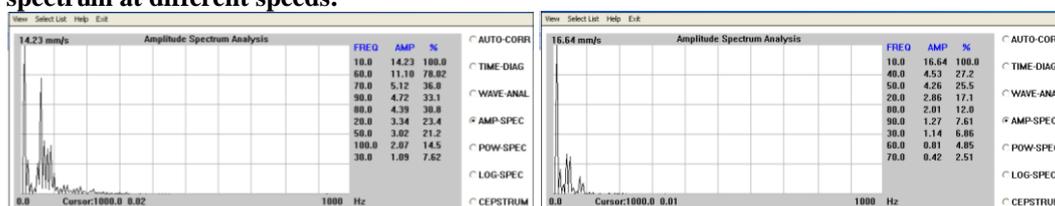
Table: 1

speed (rpm)	MNDE			MDE			PBE		
	H	V (mm/sec)	A	H	V (mm/sec)	A	H	V (mm/sec)	A
600	4.53	3.68	6.25	4.74	6.89	3.19	14.23	16.64	2.57
800	16.44	12.95	8.45	15.67	13.36	13.79	45.68	67.15	14.56
1000	76.01	80.53	58.49	22.37	22.91	59.98	83.72	158.5	50.78
1250	83.40	64.15	44.82	92.50	83.53	70.27	111.7	51.72	92.44
1440	43.17	31.15	104.93	72.09	57.90	126.1	383.8	89.06	77.61

OBSERVATIONS:

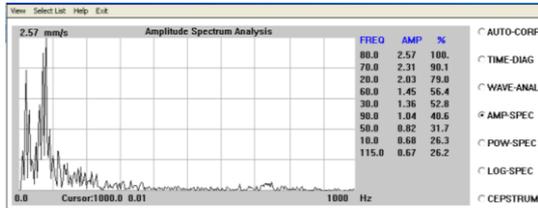
Above readings show that high vibrations are present at PBE. At all speeds vibrations in radial direction are higher than axial direction. High radial vibrations are present due to unbalance, misalignment and bending of the shaft. To determine the cause of high vibrations spectrum analysis was carried. Following are the spectrums taken at different speeds using FFT Analyzer with Mcme2.0 software.

Unbalanced spectrum at different speeds:

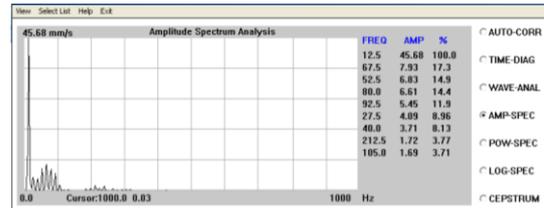


Spectrum in PBE HOZ direction at 600 rpm

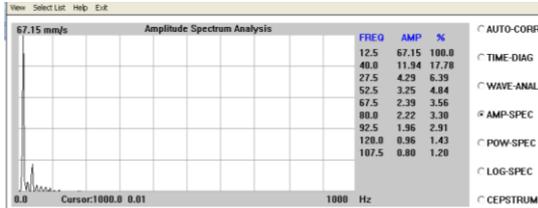
Spectrum in PBE VER direction at 600rpm



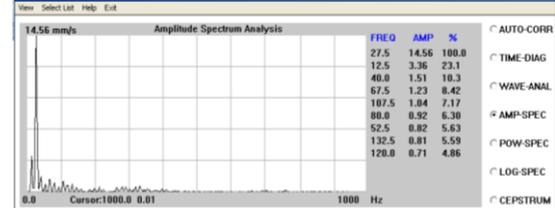
Spectrum in PBE AXL direction at 600rpm



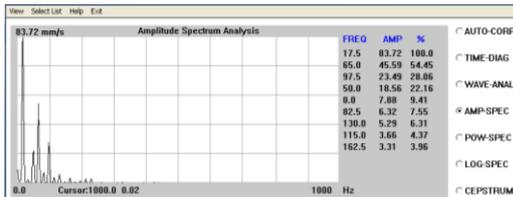
Spectrum in PBE HOZ direction at 800rpm



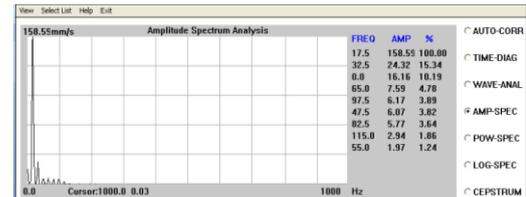
Spectrum in PBE VER direction at 800rpm



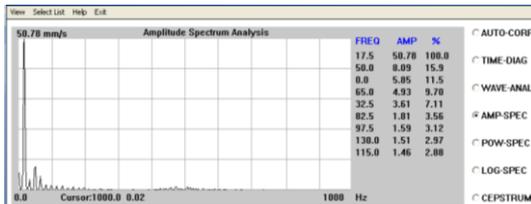
Spectrum in PBE AXL direction at 800rpm



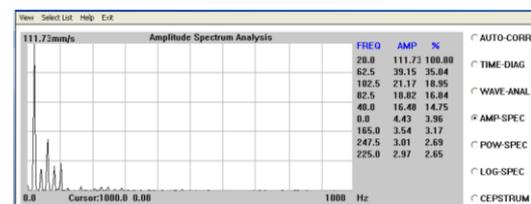
Spectrum in PBE HOZ direction at 1000rpm



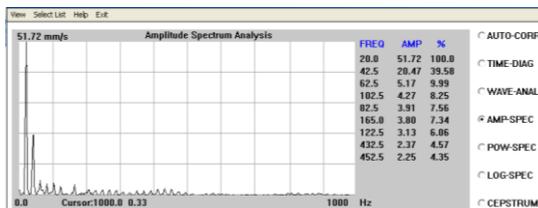
Spectrum in PBE VER direction at 1000rpm



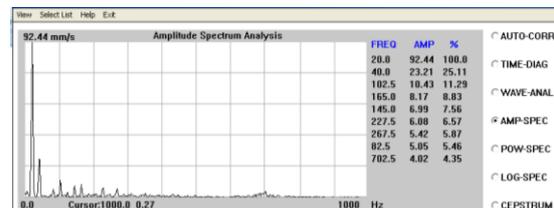
Spectrum in PBE AXL direction at 1000rpm



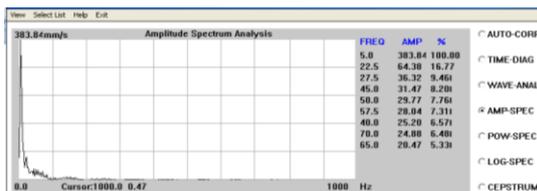
Spectrum in PBE HOZ direction at 1250rpm



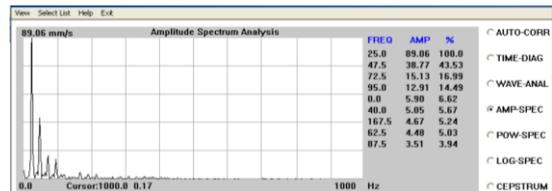
Spectrum in PBE VER direction at 1250rpm



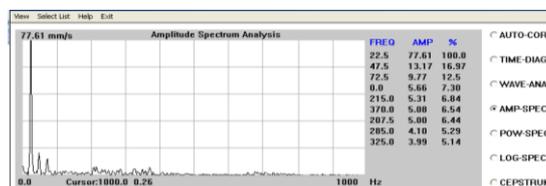
Spectrum in PBE AXL direction at 1250rpm



Spectrum in PBE HOZ direction at 1440rpm



Spectrum in PBE VER direction at 1440rpm



Spectrum in PBE AXL direction at 1440rpm

Table: 2
The following phase readings are also observed:

SPEED (rpm)	DISPLACEMENT(μm)							PHASE ANGLE IN DEGREES
	0°	30°	60°	At 90°	120°	150°	180°	
600	307	324	395	549	476	384	310	90 ^o
800	907	1028	1245	1980	1270	1372	1260	90 ^o
1000	1008	862	1604	2024	1345	1002	1119	60 ^o
1250	796	876	683	603	1311	1001	703	30 ^o
1440	735	1061	897	556	901	925	687	60 ^o

OBSERVATION FROM SPECTRUMS:

1. In all the spectrums of pillow block end 1X and its harmonics are present.
2. No bearing defective frequency peaks are present in the pillow block spectrums.
3. In the pillow block at 1000 rpm vertical spectrum “1X” is predominant and having its highest value is 158.5mm/sec.
4. In the pillow block at 1250 horizontal spectrum 1x is predominant and having the highest value is 111.7mm/sec.
5. High “1X” amplitudes in PBE horizontal and vertical spectrum indicates abnormal condition such as unbalance misalignment, looseness or resonance condition.

To determine the exact cause of high vibrations phase readings are taken. phase readings are shown in table.

Observation from Phase Analysis

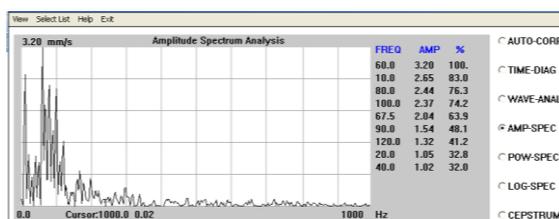
1. From the phase reading at the PBE it is observed that there is 90° phase difference between the PBE horizontal and vertical.
2. 90° phase difference between horizontal and vertical reveals that there presents an unbalance at PBE.
3. From the spectrum analysis and phase analysis it is conformed that there is a presence of mass unbalance in the PBE.

Rotor was removed and balanced weight is added diametrically opposite to unbalanced mass and vibration readings were taken after balancing. Vibration readings are shown in following table.

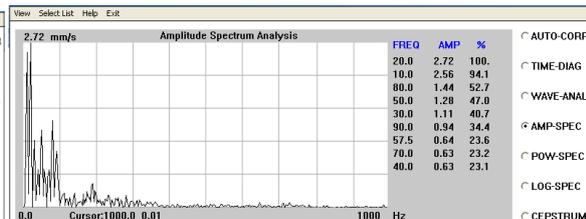
Table: 3
Vibration readings after balanced condition:

speed (rpm)	MNDE			MDE			PBE		
	H	V (mm/sec)	A	H	V (mm/sec)	A	H	V (mm/sec)	A
600	1.45	0.98	1.69	1.59	2.48	2.70	3.20	2.72	1.55
800	7.92	3.24	5.90	7.26	7.00	7.08	15.16	20.77	4.38
1000	12.36	4.43	14.89	12.15	10.37	11.56	22.95	17.85	8.01
1250	6.04	4.31	18.97	9.83	18.76	21.85	9.72	12.22	7.66
1440	10.44	25.79	21.11	6.46	12.27	20.96	13.16	20.90	9.82

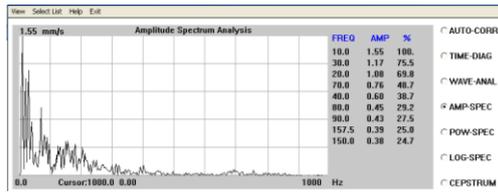
Balanced spectrum at different speeds:



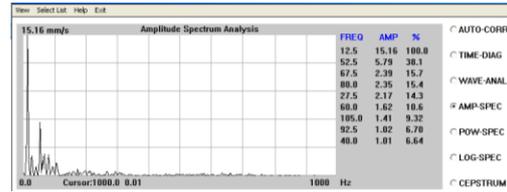
Spectrum in PBE HOZ direction at 600rpm



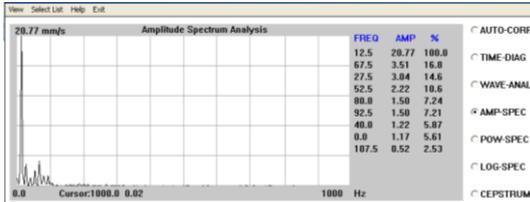
Spectrum in PBE VER direction at 600rpm



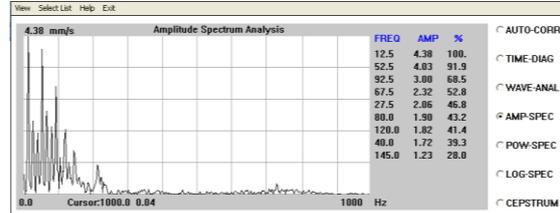
Spectrum in PBE AXL direction at 600rpm



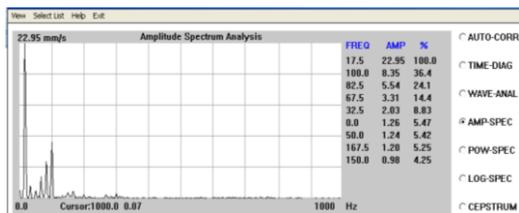
Spectrum in PBE HOZ direction at 800rpm



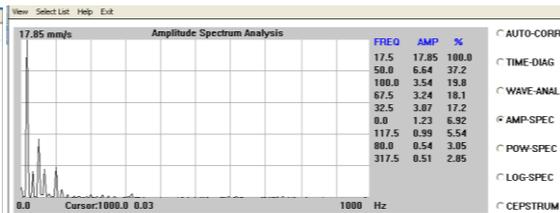
Spectrum in PBE VER direction at 800rpm



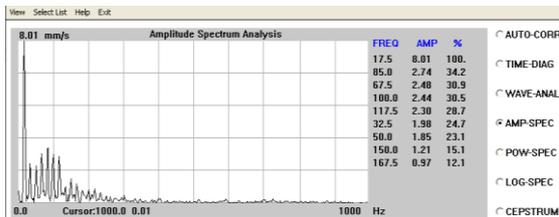
Spectrum in PBE AXL direction at 800rpm



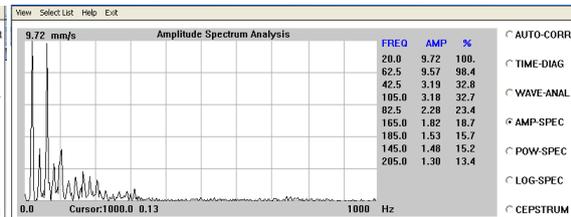
Spectrum in PBE HOZ direction at 1000rpm



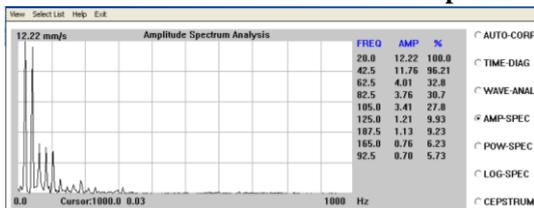
Spectrum in PBE VER direction at 1000rpm



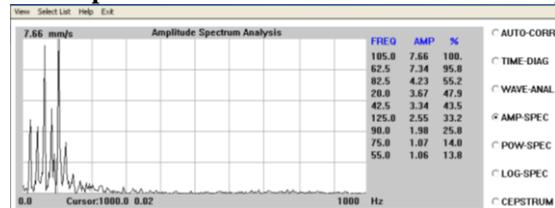
Spectrum in PBE AXL direction at 1000rpm



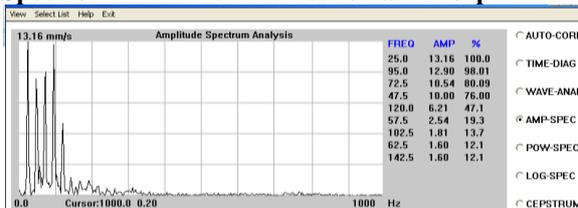
Spectrum in PBE HOZ direction at 1250rpm



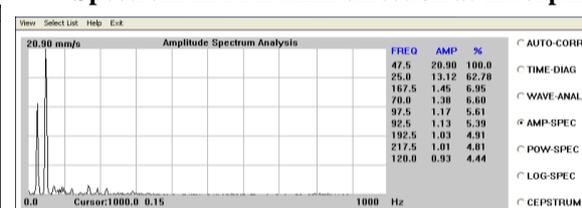
Spectrum in PBE VER direction at 1250rpm



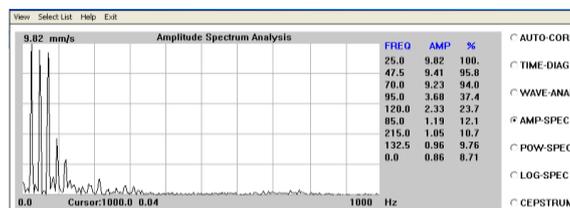
Spectrum in PBE AXL direction at 1250rpm



Spectrum in PBE HOZ direction at 1440rpm



Spectrum in PBE VER direction at 1440rpm



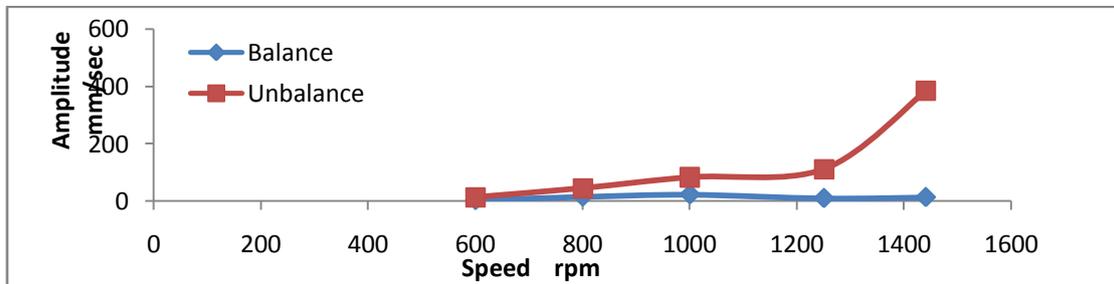
Spectrum in PBE AXL direction at 1440rpm

All vibration readings show that the readings were reduced to normal level.

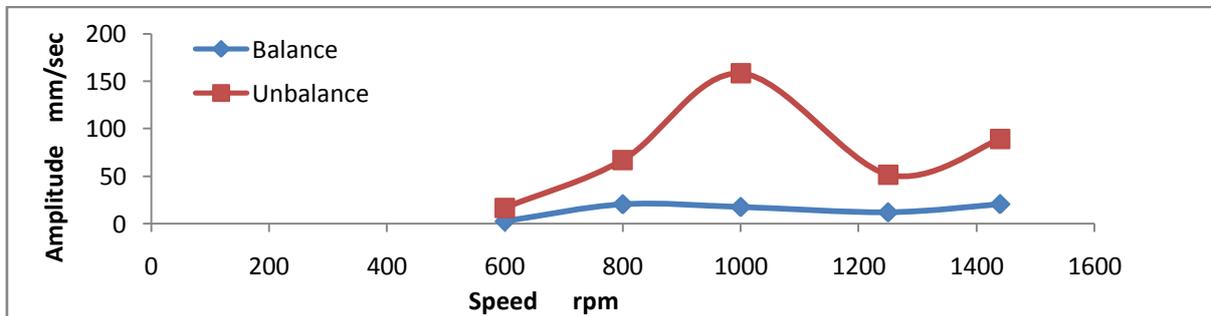
Table: 4
Comparisons of vibration amplitudes for unbalanced and balanced signals:

Speed (rpm)	Direction	MNDE		MDE		PBE	
		Unbalance	Balance	Unbalance	Balance	Unbalance	Balance
		v (mm/sec)		v (mm/sec)		v (mm/sec)	
600	H	4.53	1.45	4.74	1.59	14.23	3.20
	V	3.68	0.98	6.89	2.48	16.64	2.72
	A	6.25	1.69	3.19	2.70	2.57	1.55
800	H	16.44	7.92	15.67	7.26	45.68	15.16
	V	12.95	3.24	13.36	7.00	67.15	20.77
	A	8.45	5.90	13.79	7.08	14.56	4.38
1000	H	76.01	12.36	22.37	12.15	83.72	22.95
	V	80.53	4.43	22.91	10.37	158.5	17.85
	A	58.49	14.89	59.98	11.56	50.78	8.01
1250	H	83.40	6.04	92.50	9.83	111.7	9.72
	V	64.15	4.31	83.53	18.76	51.72	12.22
	A	44.82	18.97	70.27	21.85	92.44	7.66
1440	H	43.17	10.44	72.09	6.46	383.8	13.16
	V	31.15	25.79	57.90	12.27	89.06	20.90
	A	104.93	21.11	126.1	20.96	77.61	9.82

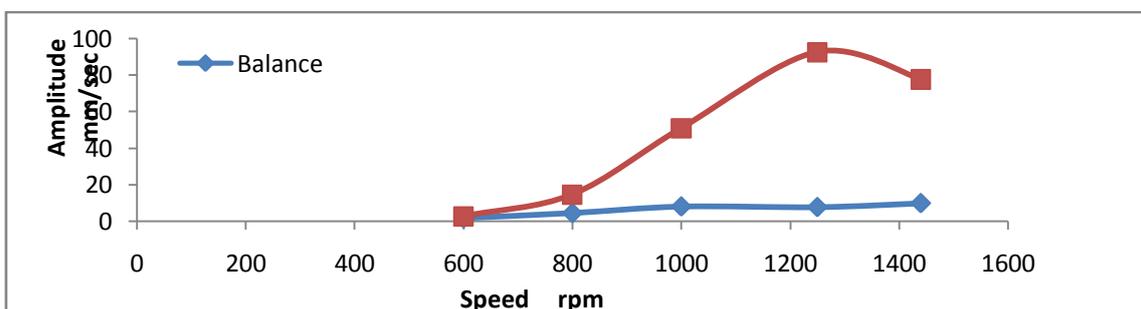
Graph shows for amplitude against speed, comparison with unbalance and balance at PBE:



Vibration amplitude at PBE HOZ against speed



Vibration amplitude at PBE VER against speed



Vibration amplitude at PBE AXL against speed

CONCLUSION

As the speed increases the amplitude at 1X is also increases for the same unbalance weight. This increase in amplitude value is because of the different unbalanced force.

Since the system frequency is nearer to 1000rpm due to the presence of resonance at this speed higher amplitudes were presented.

Phase analysis and spectrum analysis show that there presents an unbalance in the rotor.

Rotor is balanced and vibration readings are taken after balancing. It was shown that amplitude of vibration is reduced drastically.

This is an NDT method to detect the fault in rotating machine. Hence Vibration monitoring method reduces the maintenance cost when it is applied to industries and improves the profit

REFERENCE

- [1] Eshleman, R. And A. Eubanks, 1969. "On the critical speeds of a continuous rotor", J. Engineering for Industry, 91: 1180-1188.
- [2] Mitchell, L. And D.D.M. Mellen, 1995. "Torsional experimental lateral coupling in a geared high-speed rotor system", ASME Design Engineering Technical Conferences 3 (Part B), 84(2): 977-989.
- [3] D. E. Bently, *Fundamentals of Rotating Machinery Diagnostics*, Bently Pressurized Bearing Press, Minden, La, USA, 2002
- [4] Gupta, K.D. Gupta and K. Athre, 1993. "Unbalance response of a dual rotor system: theory and experiment", Transactions J. Vibration and Acoustics, 115: 427-435.
- [5] Rao, J.S., 1996. "Rotor Dynamics, 3 edition, New Age International Publishers, India.
- [6] Rao, T.N. Shiao and J.R. Chang, 1998. "Theoretical analysis of lateral response due to torsional excitation of geared rotor", Mechanism and Machine Theory, 33(6): 761-783. Coupled lateral-torsional vibration of rotor.
- [7] K.P. Ramachandran, "Vibration signature analysis for machine health monitoring and fault diagnosis," *Caledonian Journal of Engineering*, pp. 26-39, 2004.
- [8] R.K. Biswas "vibration based condition monitoring of rotating machines" national conference on condition monitoring [NCCM-2006] December 2006 pg no 34-40.
- [9] R.A. Collacat "condition monitoring" published by M c Graw Hill, New Delhi 1998
- [10] A.V. Barkov, N.A. Barkova, and A.Yu. Azovtsev, "Condition Monitoring and Diagnostics of Rotating Machines Using Vibration", VAST, Inc., St. Petersburg, Russia, 1997
- [11] Eshleman R I "Some recent advances in rotor dynamics" 3rd International conference on vibrations of rotating machinery. University of York 1984.
- [12] M.R.S. Satyanarayana et al, "Detection of gear ratio and current consumption using Motor current signature analysis", International Journal of Mechanical Engineering and Technology (IJMET), Volume 3, Issue 2, May-August (2012), pp. 643-652
- [13] M.R.S. Satyanarayana "International conference on condition monitoring On "Rotor Dynamic Considerations in Refurbishing Turbo machinery", 24-25 feb, 2011, Visakhapatnam.

Design and Optimization of Drive Shaft with Composite Materials

R. P. Kumar Rompicharla¹, Dr. K. Rambabu²
Sir C.R.R college of Engineering Department of Mechanical Eluru, India

Abstract: Automotive drive Shaft is a very important components of vehicle. The overall objective of this paper is to design and analyze a composite drive shaft for power transmission. Substituting composite structures for conventional metallic structures has many advantages because of higher specific stiffness and strength of composite materials. This work deals with the replacement of conventional two-piece steel drive shafts with a Composite material's. In this work Kevlar /Epoxy is used as composite material The design parameters were optimized with the objective of minimizing the weight of composite drive shaft. The design optimization also showed significant potential improvement in the performance of drive shaft. In this present work an attempt has been to estimate the deflection, stresses, natural frequencies under subjected loads using FEA. Further comparison carried out for both steel and composite materials and weight of the shaft is optimized and stress intensity factor found for both Steel and composite drive shafts.

Keywords: Stress intensity Factor, Deformation, Torsional stress, Drive Shaft, Modelanalysis

I. INTRODUCTION

A driveshaft is a rotating shaft that transmits power from the engine to the differential gear of a rear wheel drive vehicles Driveshaft must operate through constantly changing angles between the transmission and axle. High quality steel (Steel SM45) is a common material for construction. Steel drive shafts are usually manufactured in two pieces to increase the fundamental bending natural frequency because the bending natural frequency of a shaft is inversely proportional to the square of beam length and proportional to the square root of specific modulus. The two piece steel drive shaft consists of three universal joints, a center supporting bearing and a bracket, which increase the total weight of a vehicle. Power transmission can be improved through the reduction of inertial mass and light weight. Substituting composite structures for conventional metallic structures has many advantages because of higher specific stiffness and higher specific strength of composite materials. Composite materials can be tailored to efficiently meet the design requirements of strength, stiffness and composite drive shafts weight less than steel or aluminum of similar strength. It is possible to manufacture one piece of composite. Drive shaft to eliminate all of the assembly connecting two piece steel drive shaft. Also, composite materials typically have a lower modulus of elasticity. As a result, when torque peaks occur in the driveline, the driveshaft can act as a shock absorber and decrease stress on part of the drive train extending life. Many researchers have been investigated about hybrid drive shafts and joining methods of the hybrid shafts to the yokes of universal

joints. But this study provides the analysis of the design in many aspects.

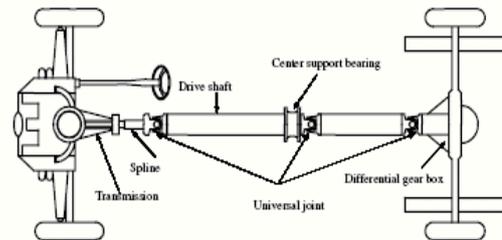


Fig 1: Schematic arrangement of Underbody of an Automobile

II. Design of composite drive shaft

2.1. Specification of the problem

The fundamental natural bending frequency for passenger's cars, small trucks and vans of the propeller shaft should be higher than 2,400 rpm to avoid whirling vibration and the torque transmission capability of the drive shaft should be larger than 154 Nm. The drive shaft outer diameter should not exceed 100 mm due to space limitations.

The torque transmission capability of the drive shaft is taken as 151 N.m the length and the outer diameter here are considered as 1.5 meters and outer diameter of the shaft is 0.072, respectively. The drive shaft of transmission system was designed optimally to meet the specified design requirements.

2.2. Assumptions

The shaft rotates at a constant speed about its longitudinal axis. The shaft has a uniform, circular cross section. The shaft is perfectly balanced, all damping and nonlinear effects are excluded. The stress-strain relationship for composite material is linear and elastic; hence, Hook's law is applicable for composite materials. Since lamina is thin and no out-of-plane loads are applied, it is considered as under the plane stress.

2.3. Merits of Composite Drive Shaft

1. They have high specific modulus and strength.
2. Reduced weight.
3. Due to the weight reduction, fuel consumption will be reduced.
4. They have high damping capacity hence they produce less vibration and noise.
5. They have good corrosion resistance.
6. Greater torque capacity than steel or aluminum shaft.
7. Longer fatigue life than steel or aluminum shaft.

2.4. Theoretical and ansys results simulation

The drive shaft for simplicity has been first idealized as a hollow cylindrical shaft which is fixed at one end and on other end which a torque of 151Nm is applied as represented below

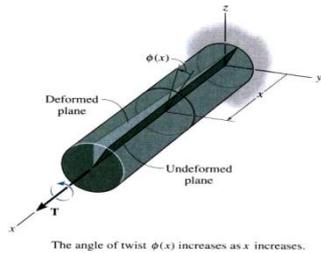


Fig 2: Shaft with torsional load

For the the hallow shaft, let
 $R_o = 0.036m$; $R_i = 0.011m$; $l = 1.5 m$; $E= 207e9$; Torque = 151Nm

Where R_o -Outer Radius of shaft
 R_i - Inner Radius of shaft
 L = Length of the shaft
 E = Young's modulus of steel (SM45C)
 T =Applied torque

Then:
 Deflection = $Y_{max} = \frac{ML^2}{2EI} = \frac{151 \times (1.5)^2}{2 \times (207e9) \times (1.178e^6)}$
 $= 0.00069 m$
 $= 0.69mm;$

Maximum deflection = $(T \times (d_o/2)) / I$

$$= \frac{151 \times (0.036)}{(\pi/2) \times [(0.036^4 - 0.011^4)]}$$

$$= 66.50 \text{ Mpa}$$

Maximum shear stress = $(T \times R_o) / J$
 $= 20.78 \text{ Mpa}$

III. 3.Simulated results for Hollow shaft in an sys

3.1. Deformation results

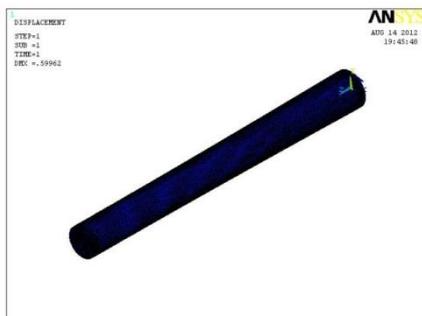


Fig 3: deformation result of steel shaft

It observed from above analysis results deformation value for steel shaft is 0.59mm.

3.2.Shear stress values

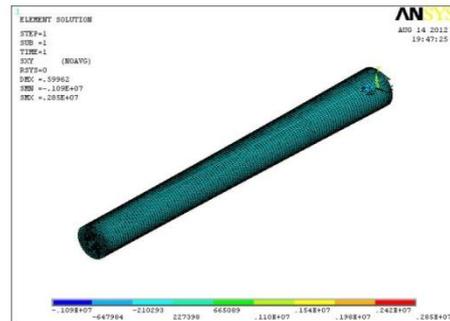


Fig 4: Shear stress value of steel shaft

It observed from above analysis results Shear stress value for steel shaft is 28Mpa

3.3. Von-Mises stress

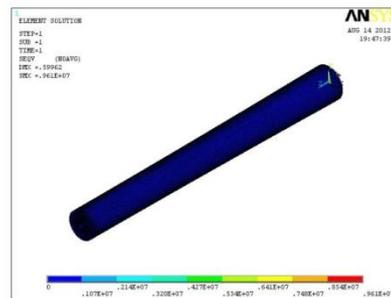


Fig 5: von-mises results

It observed from above analysis results von-Misses value for steel shaft is 96Mpa

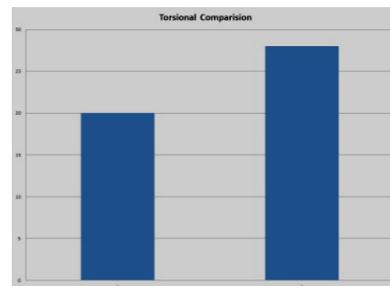


Fig 6: Torsional load comparison

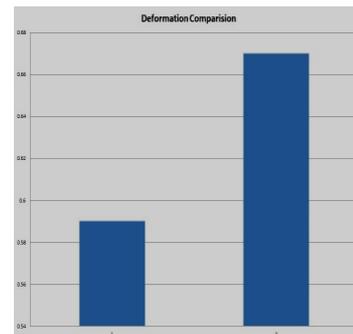


Fig 7: Deformation comparison

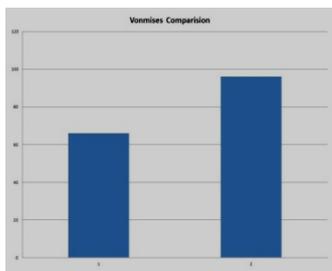


Fig 8: Von-mises Stress comparison

By comparing the theoretical values and hollow shaft analysis values it is observed that the calculated deformation value is 0.69 mm and the simulated value for deformation is .599 mm, Shear stress value calculated is 20.78Mpa for simulated it was 28Mpa, And for von-misses those values are 66Mpa and 96Mpa these results shows variation between theoretical and simulated up to 5.4 % only

IV. 4. Modeling and simulation

In this section the 3D CAD models and 3D FE Models along with the loads and boundary conditions will be presented.

Step1: 3D CATIA Model Creation was done based on considered Specifications and design consideration from Toyota Qualis specifications.

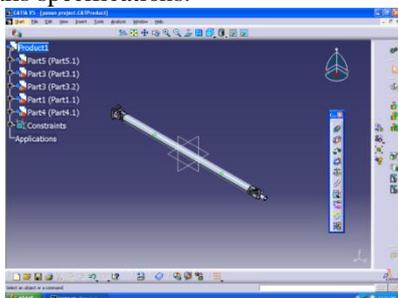


Fig 9: Catia Model

Step2: 3D FE Model Creation The 3D FE model for drive shaft was created by using FE modeling software HYPERMESH v10.0. The mesh has been generated using 2nd order Hexa elements (SOLID 95 and Solid 186) in Hypermesh.

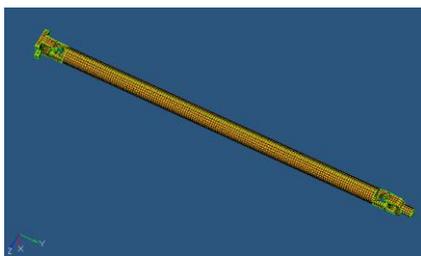


Fig 10: Hypermesh model with brick (solid 95 with contact elements)

Step-3: using above hypermesh model with boundary conditions in ansys12.0 required results are predicted.

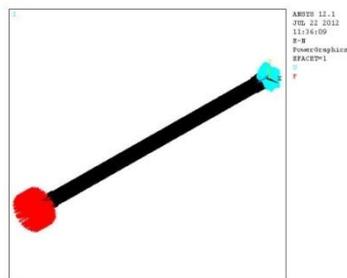


Fig 11 Ansys Model with boundary conditions

Step-4;By applying boundary conditions and loading conditions obtained results will compared and suitable material suggested which gives less torsional value and frequency nearer to steel.

4.1. Finding stress intensity value

Being able to determine the rate of crack growth, an engineer can schedule inspection accordingly and repair or replace the part before failure happens. Being able to predict the path of a crack helps a designer to incorporate adequate geometric tolerance in structural design to increase the part life. The methodology used to investigate the mechanics of crack propagation consists of the following steps:

Step 1: Introducing crack with 1mm width and 3mm depth in Catia geometric model

Step 2: Creating 3D FE model by using Hpermesh and creating fine mesh at crack located area. Using contact elements at universal joint locations

Step3: Applying Boundary conditions and to solve to get shear stress value at different locations nearer to cracktip

Step 4: Using above predicted values to plot graphs for finding stress intensity values for both Steel and Composite shafts

Step 5: Interpretation of results for both Steel and composite Intensity values

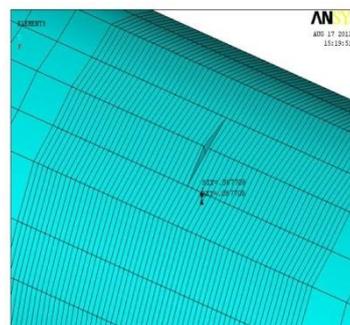


Fig 12: Shaft with crack

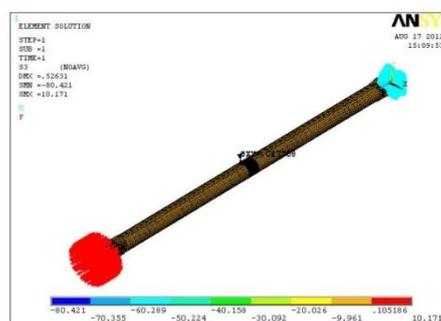


Fig13: Torsional Analysis with crack

4.2. Elements used for Analysis and its characteristics

S . No	Generic element type name	Ansys Name	Description
1	20 Node Quadratic Hexahedron	Solid 95	20 Node structural solid
2	20 Node Quadratic Hexahedron	Solid 186	20 Node structural solid
3	Quadratic Quadrilateral Contact	Conta 174	3D 8 Node surface to surface contact

4.3. Material properties used for analysis are listed below

S L no	Property	Steel (SM 45C)	Kevlar/Epoxy	units
1	Young's Modulus X direction (E_{11})	2.07e11	95.71e9	pa
2	Young's Modulus Y direction (E_{23})	-	10.45e9	pa
3	Young's Modulus Z direction (E_{31})	-	10.45e9	pa
4	Major Poisson's Ratio XY (ν)	0.3	0.34	
5	Major Poisson's Ratio YZ (ν)	-	0.37	
6	Major Poisson's Ratio XZ (ν)	-	0.34	
7	Shear Modulus XY (G_{12})	-	25.08e9	pa
8	Shear Modulus YZ (G_{23})	-	25.08e9	pa
9	Shear Modulus XZ (G_{31})	-	25.08e9	pa
10	Density	7600	1402	Kg/m ³

V. Analysis Results

Steel and Kevlar/Epoxy shaft deformation comparison.

5.1. Steel drive shaft deformation result



Fig 14: Steel shaft deformation results



Fig 15: Kevlar/Epoxy drive shaft deformation results

By considering above results it is observed that steel shaft having deformation value of 0.589 mm and Kevlar/Epoxy drive shaft having deformation value of 8.1 mm

5.2. Torsional stress comparison

It is observed from below anlysis results steel shaft having maximum stress value in of 53.80Mpa XY direction and Kevelar /Epoxy shaft having maximum shear stress value in XY direction is 49.82Mpa only.The ansys simulated values are as shown below



Fig 16: Steel shaft torsional analysis



Fig 17: Kevlar/Epoxy torsional analysis value

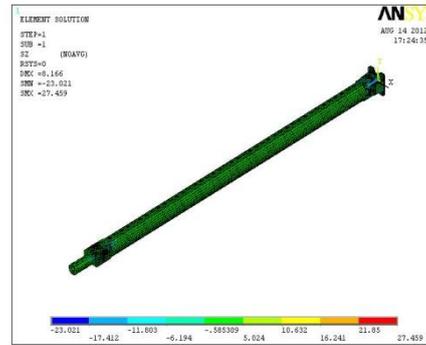


Fig 21: Kevlar/Epoxy buckling stress values

5.3. Model analysis results



Fig18: Steel shaft Model analysis



Fig19: Model analysis of Kevlar/Epoxy

It is observed from above model analysis results the natural frequency of steel shaft is 3.7Hzs and 2.78Hzs for Kevlar /Epoxy so it is from predicted values it is observed natural frequency values are very nearer to each steel and Kevlar/epoxy shafts.

5.4. Eigen buckling Result's (inz direction)

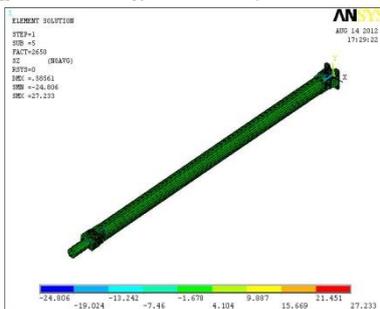


Fig 20: Steel shaft buckling

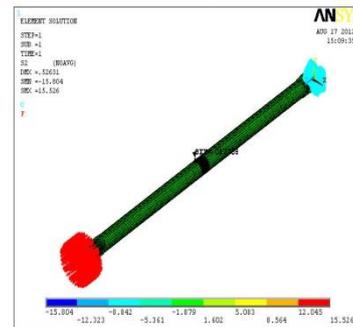


Fig 22: Torsional analysis of shaft with crack

5.6. Steel shaft with cut section with stresses at crack tip

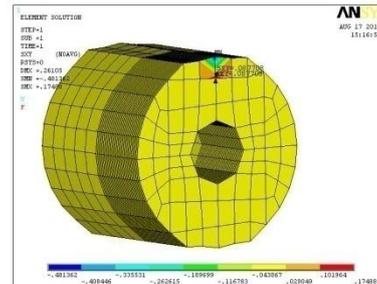


Fig 23: Steel shaft with crack tip cross-section

5.7. Steel shaft predicted intensity Values

S.N	Distenc e (r)	Shear stress value in XY direction(σ)	K_I value ($\sigma\sqrt{(r)}$)
1	29.51	0.08770	0.472
2	28.93	0.00904	0.09
3	30.59	-0.000055	-0.0003

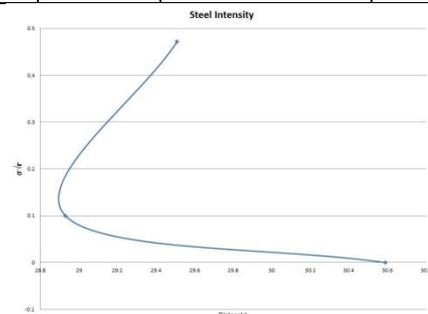


Fig 24 : Steel Shaft intensity Graph

By considering graph plotted between Distance (r) and stress $\sigma\sqrt{r}$ from crack tip the stress intensity factor K_{III} value for steel shaft is observed as 0.13Mpa \sqrt{mm} .

5.8. Kevlar/Epoxy shaft with cut section with stresses at crack tip

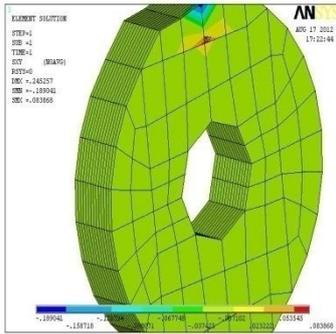


Fig 25: Composite shaft with crack tip cross-section

5.9. Composite shaft predicted intensity values

S.NO	Distance (r)	Shear stress value in XY direction(σ)	K_I value ($\sigma\sqrt{r}$)
1	29.51	0.000741	0.003993
2	28.93	0.0001126	0.0007
3	30.59	-0.000056	-0.0003

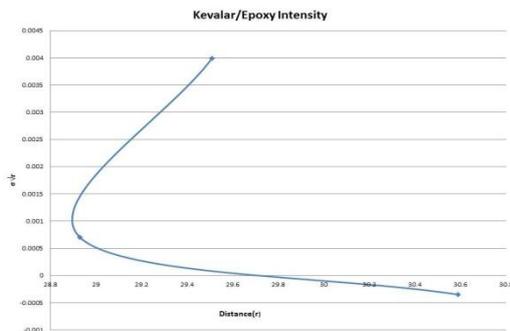


Fig 26: Composite Shaft intensity Graph

By considering graph plotted between Distance (r) and stress $\sigma\sqrt{r}$ from crack tip the stress intensity factor K_{III} value for composite shaft is observed as 0.012 Mpa \sqrt{mm}

VI. Results summary

S.No	1	2
Material	Steel	Kevlar/Epoxy
Defromation in mm	0.5816	8.16
Number of layers	-	2
Angle of ply	-	± 45
Natural Frequey in HZ	3.76	2.04
Trosional Stress value in N/mm2	53.80	49.82
Buckl-ing Stress Value inN/mm2	27.45	27.23

Weight reduction in %	-	23
-----------------------	---	----

6.1. Stress intensity values

S.No	Material	Stress Intensity value in Mpa \sqrt{mm} .
1	Steel	0.13
2	Kevlar/Epoxy	0.012

VII. Conclusion

- 1) The usage of composite material has resulted to inconsiderable amount of weight saving in the range of 28 % when compared to conventional steel shaft
- 2) Taking into considerations the weight saving, deformation, shear stress induced and resonant frequencies it is evident that Kevlar/Epoxy composite has the most encouraging properties to act as replacement for steel out of the considered two materials .
- 3) The presented work was aimed to reduce the fuel consumption of the automobile in the particular or any machine, which employs drive shafts ,in general it is achieved by using light weight composites like Kevlar/Epoxy
- 4) The presented work also deals with design optimization i.e converting two piece drive shaft (conventional steel shaft) in to single piece light weighted composite drive shaft.
- 5) The drive shaft of Toyota Qualis was chosen for determining the dimensions, which were used then used for the material properties of composites were used the stability of drive shaft is ensured by limiting the include values with in the permissible range in Ansy 12.0
- 6) The stress intensity value (K_{III}) at crack tip is observed for composite driveshaft is low.

References

Journal Papers

- [1] 73332270 Design and Analysis of a Propeller Shaft of a Toyota Qualis by "Syed Hasan"
- [2] Mechanics of laminated composite plates and shells: theory and analysis – by Junuthula Narasimha Reddy.
- [3] Optimal Sizing and Stacking Sequence of Composite Drive shafts- Thimmigowda rangaswami,Sabapathy Vijayarangan.
- [4] Polymer Matrix composites In Drive line Applications-Drf Andrew Pollard, GKN Technology,Wolverhampton , UK.
- [5] Static Torsion Capacity of Hybrid Aluminum Glass Fiber Composite Hallow Shaft-S.A.Mutasher , B.B.Sahari and A.M.S Hamouda, S.M.Sapuan.
- [6] Automotive Composite Drive shaft: Investigation of Design variables Effects- M.A.Badie, A.Mahdi, A.R.Abotalib, E.J.Aabdullah and R.Yonus.
- [7] Static and dynamic characteristics of composite shafts- S.A.Mutasher , B.B.Sahari and A.M.S Hamouda, S.M.Sapuan.
- [8] Hurd, N.J.Torsional performance of Drives hafts for Vehicle Driveline Applications- j.of SAE 960573,1996.

- [9] Swason, S.R.(1997),Introduction to design and analysis With advanced composite materials, *Prentice Hall, Upper Saddle River, Newjersey*.
- [11] Comprehensive composite Materials- By *Anthony Kelly, Carl H, Zweben*.
- [12] "Development of new Cposites" . *Kawarada K,Yamagata H, Wakamatsu M, Sekiyama K,*
- [14] Stress intencity factores under combined loads of torsion and bending loads by *A.E.Ismail,A.K.Arrifin,S.Abdullah,M.J.Ghazali*.
- [15] Computation of Mixed-Mode Stress Intensity Factors for Cracks in Three-Dimensional Functionally Graded Solids –By *Matthew C. Walters,Glaucio H. Paulino and Robert H. Dodds*
- [16] OPTIMAL DESIGN AND ANALYSIS OF AUTOMOTIVE COMPOSITE DRIVE SHAFT by-*T.Rangaswamy, S. Vijayarangan, R.A. Chandrashekar, T.K. Venkatesh and K.Anantharaman*
- [17] Optimum Design And anlysis of Automotive Drive shaft For An Automobile by- *Gummadi Snjay,Akula Jagadeesh Kumar from Karlskrona,Swedan*.
- [18] DESIGN AND MODAL ANALYSIS OFCOMPOSITE DRIVE SHAFT FOR AUTOMOTIVE APPLICATION By- *Mohammad Reza Khoshravan,Amin Paykani,Aidin Akbarzadeh, Faculty of Mechanical Engineering, University of Tabriz ,Tabriz, Iran*
- [19] Numerical Evaluation of Stress Intensity Factors (KI) J-Integral Approach by- *Guillermo A. Riveros*
- [20] Mode III stress intensity factors for edge-cracked circular shafts, bonded wedges, bonded half planes and DCB's By- *A.R. Shahani Faculty of Mechanical Engineering, K.N.T. University of Technology, Vafadar-e-Sharghi St., 016765-3381, Tehran, Iran*
- [22] Massachusetts Institute of Technology *Department of Mechanical Engineering Cambridge, Laboratory Module No. 4 Isotropic Linear Elastic Stress Concentration*
- [23] Introduction to Fracture Mechanics David Roylance *Department of Materials Science and Engineering,Massachusetts Institute of Technology, Cambridge, MA 02139,June 14, 2001.*
- Books:**
- [1] "Design Data"- Data book of engineering.
- [2] Stress Intensity Factor and Limit Load Handbook by *S .Al Laham,Structural Integrity Branch*
- [3] Ramamurti,V.(1998), Computer-aided mechanical design and analysis ,3rd Edition, *McGraw-Hill,New York*.

Analyzing Fault Diagnosis Using PPDG & Its Application's

Sambangi Swathi (M. Tech.), Mudiganti Vijaya Bhaskar (Associate Professor)

Department of Computer Science Gokul Institute of Technology and Sciences, Bobbili, India.

ABSTRACT: In this paper a Model based Fault Localization Technique is used, which is called Probabilistic Program Dependence Graph, a contemporary model that scans the internal behavior of the project over a set of test inputs. The PPDG model captures the conditional statistical dependence and independence relationships among program elements in such a way that it facilitates by making probabilistic inferences about program behaviors. PPDG construction is enhanced by Program Dependence Graph (PDG) that represents the structural dependences of a program with estimations of statistical dependences between node states, which are computed from the test set. The acquirement of probabilistic graphical models, which are widely used in applications such as medical diagnosis are the basis for the PPDG. This paper discusses the algorithms needed for constructing PPDGs and the applications of the PPDG to fault diagnosis. This paper also outlines that Probabilistic Program Dependence Graphs can simplify fault localization and fault comprehension.

Keywords: PPDG, Fault Localization technique, dependencies, Graphs.

I. INTRODUCTION

In software engineering applications to abstract relevant relationships between program elements or states, a variety of graphical models have been used and thereby those models facilitate program analysis and understanding. These models include control-flow graphs, call graphs, finite-state automata, and program dependence graphs. If the models are generated by static analysis, they indicate that certain occurrences are possible at run time where as models produced by dynamic analysis indicate what actually does occur during one or more executions. The commonly used graphical models do not support making conclusions about the program behavior and also limits the utility of the models for reasoning about the causes and effects of inherently uncertain program behaviors, such as runtime failures.

In this paper, we show how the program dependence graph can be used to know the program behavior. The model captures the conditional statistical dependence and independence relationships among program elements in a way that facilitates making probabilistic inferences about program behaviors. We call this model a Probabilistic Program Dependence Graph (PPDG). Our technique produces the PPDG for a program by augmenting its program dependence graph automatically. The technique associates a set of abstract states with each node in the PPDG. Each abstract state represents a (possibly large) set of concrete nodes states in a way that is chosen to be relevant to one or more applications of PPDGs. Each node has a conditional probability distribution that relates the states of the node to the states of its parent nodes. The technique estimates the parameters of the probability distribution by analyzing executions of the program, which are induced by a set of test cases or captured program inputs.

II. BACKGROUND & PREVIOUS WORK

In this section, we briefly review two models that form the basis for the Probabilistic Program Dependence Graph. The first is the program dependence graph, which represents structural dependences between program

statements. The second is a dependency network, which is a type of probabilistic graphical model that represents conditional dependence and independence relationships between random variables.

A) Program Dependence Graph

Before describing a program dependence graph, we define and illustrate the control flow graph, which is used to construct the program dependence graph.

Definition 1: A control flow graph for a program P is a pair(N,E), where N is a set of nodes that represents statements in P and E is a set of directed edges in which each edge (ni, nj) represents the flow of control from node ni to node nj. Edges representing conditional branches are labeled to represent the conditions under which those edges are taken.

To illustrate, consider example program findmax, shown in Fig. 1, which outputs the maximum of a set of integers. Fig. 2 shows the control flow graph for findmax. In the graph, each node is labeled with the number of the program statement that it represents, and each edge shows the flow of control between the corresponding statements. For example, node 1 represents the first statement in the program and node 10 represents the last statement in the program. For another example, node 4 has two outgoing edges: Edge (4, 5) is taken if the condition at 4 is true (i.e., the while loop is entered) and edge (4, 10) is taken if the condition at 4 is false.

```
0 void findmax() {
1   int i = 0;
2   int n = read_int();
3   int max = 0;
4   while (i < n){
5       int v = read_int();
6       if (v > max)
7           max = v;
8       i++;
9   }
10  print(max);
11 }
```

Fig.1:An Example Program of FindMax

Definition 2: In a control flow graph G, node n1 is control dependent on node n2 if n2 has outgoing edges e1 and e2 such that 1) every path in G starting with e1 and ending with an exit node contains n1 and 2) there is a path starting with e2 and ending with an exit node that does not contain n1.

For example, in Fig. 2, nodes 1-4 and 10 are control dependent on the program entry point—by convention, a dummy edge is added from the program entry point to each program exit point so that top-level nodes are control dependent on the entry. Nodes 5, 6, and 8 are control dependent on node 4, and node 7 is control dependent on node 6.

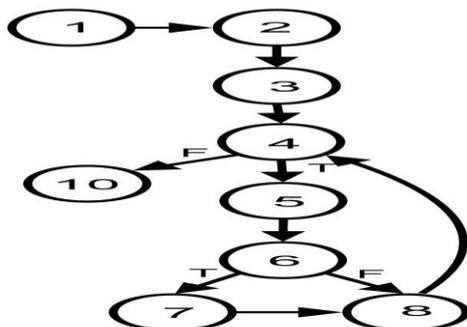


Fig.2: FindMax with its CFG

Definition 3: In a control flow graph G, node n1 is data dependent on node n2 if 1) n2 defines a variable v, 2) there is a path in G from n2 to n1 that does not redefine v, and 3) n1 uses v.

For example, in Fig. 2, nodes 4 and 8 are data dependent on node 1 for variable i and node 4 is data dependent on node 2 for variable n. Now, using control dependence and data dependence, we can define a program dependence graph [8].

Definition 4: A program dependence graph (PDG) is a directed graph whose nodes represent program statements and whose edges represent data and control dependences. Labels on the control dependence edges represent the truth values of the branch conditions for those edges, and labels on data dependence edges represent the variables whose values flow along those edges.

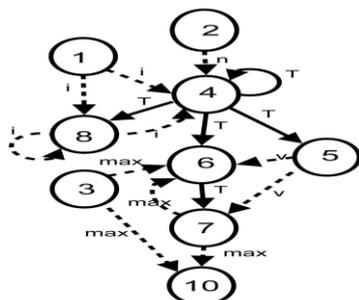


Fig.3:FindMax with its PDG

Fig. 3 shows the PDG for program findmax in Fig. 1. The nodes in the PDG are labeled with the line numbers of the corresponding statements in the program. Solid edges represent control dependences between nodes and dotted edges represent data dependences between nodes. Labels on the control dependence edges are either “T” for true or “F” for false. Labels on the data dependence edges represent the

variables involved in the data flows between the nodes. For example, in Fig. 1c, node 6 is control dependent on node 4 and it is data dependent on nodes 3, 5, and 7. The control dependence edge between node 4 and node 6 has the label “T,” which indicates that node 6 is executed when the branch condition at node 4 is true. The data dependence edge between node 3 and node 5 has the label “max,” which indicates that the value of variable “max” at node 3 flows to node 6. The precision of a PDG depends on the precision of the underlying analyses. For example, the precision of the pointer analysis affects the precision of the data dependences.

B) Dependency Network

A dependency network is a type of probabilistic graphical model.

Definition 5: A probabilistic graphical model is an annotated graph that captures the probabilistic relationships among a set of random variables. The nodes in the graph represent random variables and the edges represent conditional dependences between the random variables.

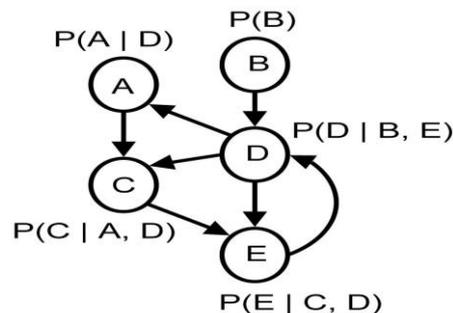


Fig.4: An Example for Dependency Network

III. SYSTEM ANALYSIS & DESIGN

A) System Analysis

A probabilistic program dependence graph (PPDG) is created by transforming the PDG of a program into a dependency network. We use a dependency network because it permits directed cycles, which are present in the PDGs of typical programs because of loops. Henceforth, we use the terms “loop” and “cycle” interchangeably. The process of producing a PPDG consists of five main steps, as illustrated in Fig. 5. First, the PDG-generation step generates the PDG of the input program P. Second, the PDG-transformation step takes the PDG, and transforms it by structurally changing the PDG and specifying states at nodes in the PDG, which results in a transformed PDG. Third, the Instrumentation step inserts probes into P to gather the execution data needed to estimate the parameters of the PPDG, and produces the instrumented program P0. Fourth, the Execution step executes P0 with its test suite TP to generate the execution data. Finally, the Learning step generates a PPDG based on the execution data and the transformed PDG by estimating the parameters of the PPDG. The resulting PPDG is formally defined as follows:

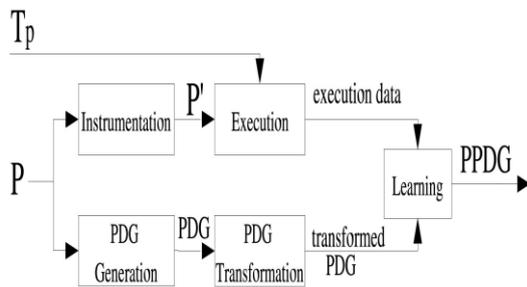


Fig.5: Construction of PPDG

B) Module Analysis

There are four modules in our paper

- **New user Registration:**

User can be providing the personal information at the registration phase. These sensitive user results of information can be placed inside the database. Automatically user can be get the information or credentials (username and password).

- **Testing, Debugging and Maintenance:**

Based on the user credentials enter inside the project deployment to enter inside the homepage. Browse or select the program and deploy the program. Program can be executed indently the relationship of information from one state to another state. Identify the program behavior of information and faults comprehension information. Whenever to identify the faults automatically to define that information like quality representation process.

- **Probabilistic Program Dependence Graph:**

PPDG produces the PPDG for a program by augmenting its program dependence graph automatically. The technique associates a set of abstract states with each node in the PPDG. Each abstract state represents a (possibly large) set of concrete nodes states in a way that is chosen to be relevant to one or more applications of PPDGs. Each node has a conditional probability distribution that relates the states of the node to the states of its parent nodes. The technique estimates the parameters of the probability distribution by analyzing executions of the program, which are induced by a set of test cases or captured program inputs.

- **Show the dependency network:**

There are different kinds of probabilistic graphical models, including Bayesian networks, Markov random fields, and dependency networks. Bayesian networks are directed acyclic graphs, whereas Markov random fields are undirected graphs. Dependency networks are similar tom Bayesian networks except that they may contain cycles.

C) System Design

Class diagrams model class structure and contents using design elements such as classes, packages and objects. The association relationship is the most common relationship in a class diagram. The association shows the relationship between instances of classes. For example, the class Order is associated with the class Customer. The multiplicity of the association denotes the number of objects that can participate in then relationship.

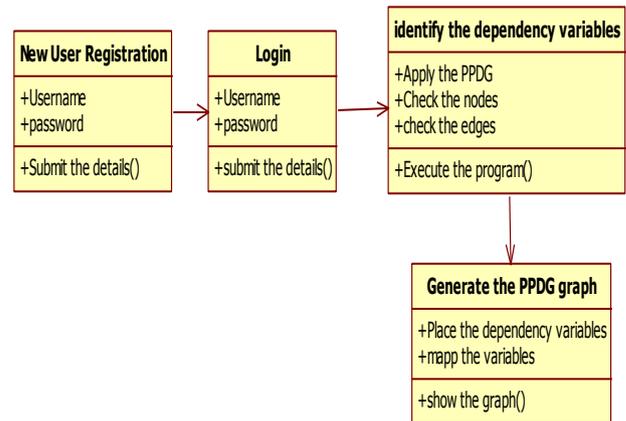


Fig.6: Inter-Operational Class Diagram for Framework

A use case illustrates a unit of functionality provided by the system. The main purpose of the use-case diagram is to help development teams visualize the functional requirements of a system, including the relationship of "actors" (human beings who will interact with the system) to essential processes, as well as the relationships among different use cases. Use-case diagrams generally show groups of use cases -- either all use cases for the complete system, or a breakout of a particular group of use cases with related functionality

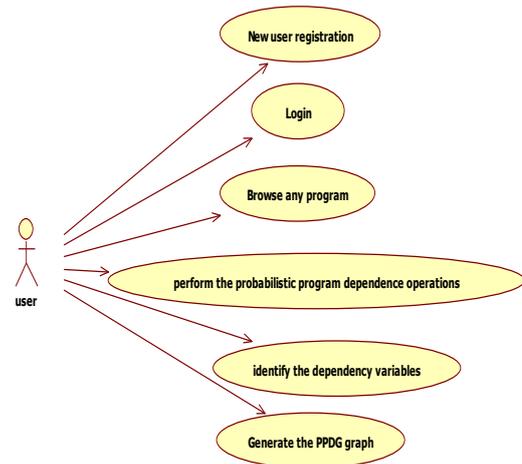


Fig.7: Inter-operational Usecase Diagram for the Framework

IV. RESULTS

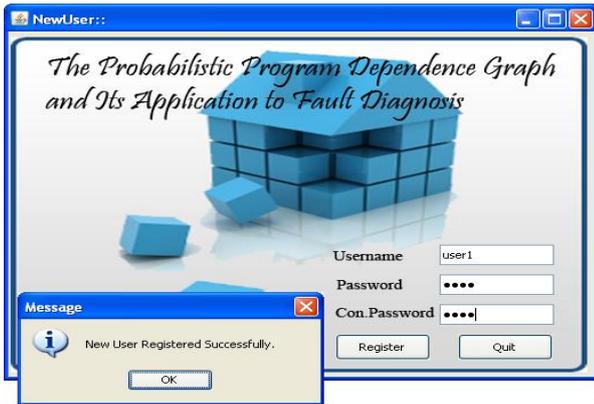


Fig.7: New User Registration

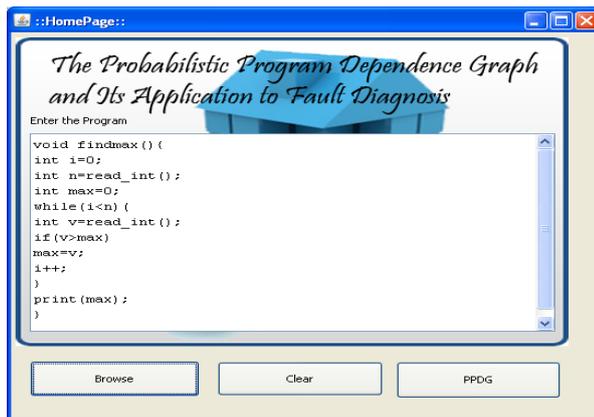


Fig.8: Giving Program to find the Dependency

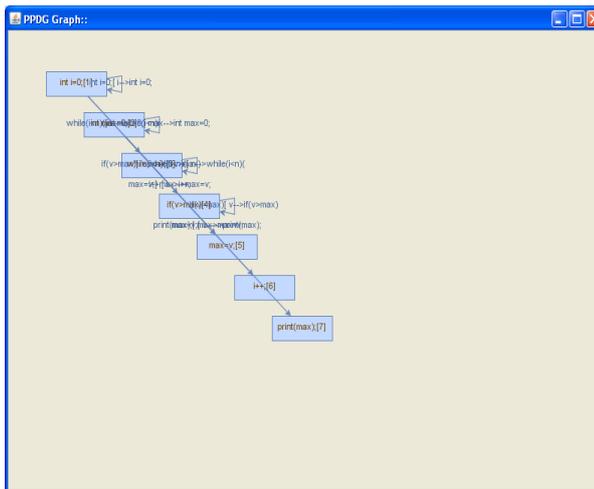


Fig.9: Extracting the PPDG Graph

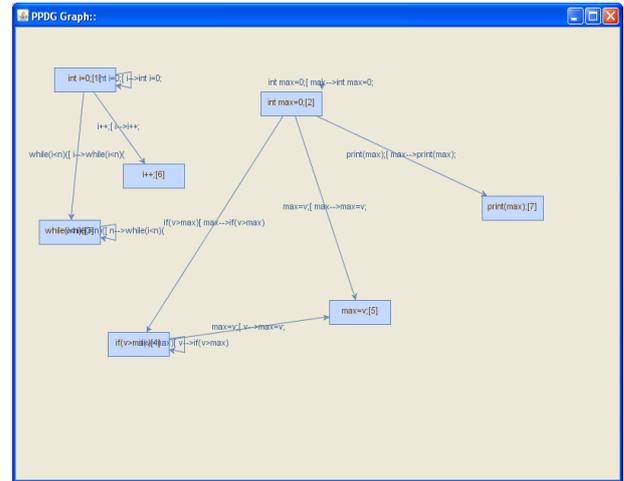


Fig.10: Identifying the PPDG Graph

V. CONCLUSION

In this paper, we presented our technique PPDG that uses the program dependence graph to create a novel probabilistic graphical model .PPDG depends on the PDG that captures the statistical dependences among program elements and enables the use of probabilistic reasoning to analyze program behaviors. This paper has discussed the two applications of the PPDG for Software engineering tasks. For the first task fault localization it has shown that how the PPDG can be used to overcome the limitations of current fault-localization techniques by introducing a simple ranking –based algorithm. Fault comprehension as a second task of application, we presented an algorithm that exploit the interpretive nature of the PPDG. *RankCP* is an algorithm which uses the PPDG to rank statements to assist in fault localization and *FaultComp*, which uses the PPDG to generate explanations to aid in fault comprehension. *RankCP* and *FaultComp* were implemented for the evolution of the PPDG.

The most critical part of our PPDG construction is the execution information, which is used to estimate the parameters of the PPDG. This execution information is dependent on the test suite that is executed by the instrumented program. Our experiment, although limited, suggests that our technique is more efficient than existing techniques that consider single failing executions. We used PPDG in this paper depending on PDG, and hence the statistical dependences across the functions were not captured by PPDG. In the future we will use the PPDG as the base on the interprocedural PDG and facilitate PPDG in such a way that it can capture the statistical dependencies of program elements.

We have shown the potential utility of applying the PPDG to the problem of fault diagnosis but we believe that it has other applications. We therefore plan to investigate the potential application of the PPDG to other software engineering tasks.

REFERENCES

- [1] S. Bates and S. Horwitz, "Incremental Program Testing Using Program Dependence Graphs," Proc. Symp. Principles of Programming Languages, pp. 384-396, Jan. 1993.
- [2] J.F. Bowering, J.M. Rehg, and M.J. Harrold, "Active Learning for Automatic Classification of Software Behavior," Proc. Int'l Symp. Software Testing and Analysis, pp. 195-205, July 2004.
- [3] S. Thrun, "Robotic Mapping: A Survey," Exploring Artificial Intelligence in the New Millennium, pp. 1-35, Morgan Kaufmann Publishers, Inc., 2002.
- [4] W. Weimer and G. Necula, "Mining Temporal Specifications for Error Detection," Proc. Int'l Conf. Tools and Algorithms for the Construction and Analysis of Systems, pp. 461-476, Apr. 2005.
- [5] M. Weiser, "Program Slicing," Proc. Int'l Conf. Software Eng., pp. 439-449, Mar. 1981.
- [6] J. Ferrante, K.J. Ottenstein, and J.D. Warren, "The Program Dependence Graph and Its Use in Optimization," ACM Trans. Programming Languages and Systems, vol. 9, no. 3, pp. 319-349, July 1987.
- [7] S. Galan, F. Aguado, F.J. Diez, and J. Mira, "NasoNet, Joining Bayesian Networks, and Time to Model Nasopharyngeal Cancer Spread," Artificial Intelligence in Medicine, pp. 207-216, Springer, 2001.
- [8] R. Alur, P. Cerny, P. Madhusudan, and W. Nam, "Synthesis of Interface Specifications for Java Classes," Proc. Symp. Principles of Programming Languages, pp. 98-109, Jan. 2005.
- [9] X. Zhang, N. Gupta, and R. Gupta, "Pruning Dynamic Slices with Confidence," Proc. Conf. Programming Language Design and Implementation, pp. 169-180, June 2006.
- [10] B. Liblit, M. Naik, A.X. Zheng, A. Aiken, and M.I. Jordan, "Scalable Statistical Bug Isolation," Proc. Conf. Programming Language Design and Implementation, pp. 15-26, June 2005.
- [11] C. Liu, X. Yan, L. Fei, J. Han, and S.P. Midkiff, "SOBER: Statistical Model-Based Bug Localization," Proc. European Software Eng. Conf. and ACM SIGSOFT Symp. Foundations of Software Eng., pp. 286-295, Sept. 2005.

AUTHORS LIST



SWATHI SAMBANGI received her Bachelor's Degree in Information Technology in GMR Institute of Technology Affiliated to JNTU Kakinada and pursuing Masters of Technology in Software Engineering in Gokul Institute of Technology and Sciences affiliated to JNTU Kakinada. Her research areas of interest are Software Engineering, Computer Networks and Data Mining.

MUDIGANTI VIJAYA BHASKAR, Completed his B.Tech ,Computer Science in (A.U) 1997 and M.tech,Computer Science (A.U) 2000 .Presently Pursuing Ph.D in Soft computing in Andhra University . His area of interest are Image processing, Artificial Intelligence,Robotics and Soft computing.

Finger Based Indoor Robot by Gesture Recognition

Suneetha.G¹, A. Kabir Das², N. V. G. Prasad³.

^{1,2,3} Department Of Electronics and Communication Engineering, SITE, Tadepalligudem, India.
¹M.Tech Student, ²Assistant Professor & ³Associate Professor.

Abstract: Processing Real-Time image sequence is now possible because of advancement of technological developments in digital signal processing, wide-band communication, and high-performance VLSI. Recently, the demand for the indoor robots has increased. Therefore, increased opportunities for many people to operate the robots have emerged. However, for many people, it is often difficult to operate a robot using the conventional methods like remote control. To solve this problem, we propose a robot operation system using the hand gesture recognition. Our method pays attention to the direction and movement of the hand. We were able to recognize several gestures in real-time.

I. Introduction

Recently, the demand for the indoor robots, such as auto cleaning and transporter robots has increased. However, the indoor robots have two problems. First, it is usually not easy for many people to operate them because the operations are often difficult and confusing. Second, safety of the robots while operating in the real environment. In this work, we focused on solving the robot operation problem. In order to solve the problem, we thought it is necessary and desirable to construct an easy operation system for everyone. Therefore, we proposed a robot operation using hand gestures. This is because hand gestures are easy to understand and intuitive for many people. We recognize the hand gesture divided into motion and hand direction. We used optical flow for hand motion recognition and Harris operator for hand direction recognition.

Related works include also using the hand gesture to operate a robot. This work uses camera and ultrasonic wave sensors to recognize hand gesture. However, our work uses only a camera for easy

Application and to construction of a simple system. Moreover, in, an input interface for the computer by the hand gesture is proposed. Our work estimates hand motion by finger position. However, we need to recognize gestures even when we cannot extract a large hand area in an image. In this work, we propose a method to recognize gestures at a distance of 1 to 2 meters from the robot.

II. Environment

2.1 System Requirements

In constructing an indoor robot, we thought that it should meet the following requirements:

1. It can operate in real time.
2. User friendly.

Therefore, we have developed a robot system as shown in Fig.1. Autonomous robot operation is realized by computer image processing results. High-speed image processing becomes possible because the main processing is done by a computer. We can then send the control commands from the computer to the robot.

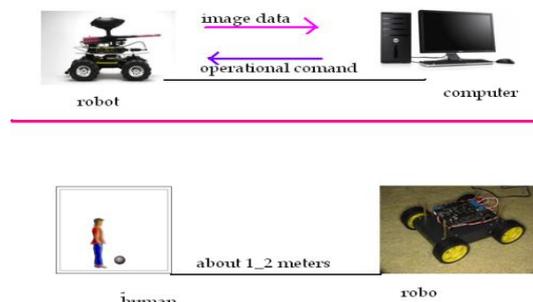


Fig. 1: Robot System

2.2 Camera and Computer Spec

In this section, we describe the specifications of the camera and computer. Camera can obtain the images at 640x480 pixels and at 30 frames per second. The video format is Motion JPEG. In this work, we use 320x240pixels image at 15 frames per second. The specifications of computer used for image processing are: The CPU of the is Core2Duo (3.0GHz) processor and the memory is 2[GB].

2.3 Definition of Gesture

In this section, describe the gestures used this paper. We define the four gestures to direct the operation of the robot. These are "Go straight", "Stop", "Turn Left" and "Turn Right". The starting home position is a shown in Figure 2. The gestures are shown in Figures 3, 4, 5 and 6 respectively.



Fig.2 Home position



Fig.3 Down



Fig.4 Stop



Fig.5 Right



Fig.6 Left

Using these gestures, the following two operations can be defined.

1. Which direction the hand moved
2. The final direction of the hand

This work proceeds by setting these two conditions. In the following sections, we will explain how to achieve these conditions.

III. Proposed Method

3.1 Flow of Method

The flow of the proposed method is shown in Figure 7. This work consists of two stages, hand motion detection and final direction estimation.

At first, skin color regions are extracted from the input image. The moving parts are considered the hands. Next, on the extracted hand areas, features are extracted and using the features, the hand final direction is detected. Then, based on the final direction of the hand, the gesture is recognized. The details are as follows.

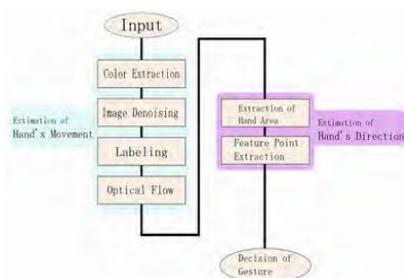


Fig.7 Flow Chart

3.2 Estimation of Hand's Movement

3.2.1 Color Extraction

To detect the human hand, we extract skin color area. The raw data provided by the camera is in the RGB format. However, detecting skin color is difficult using the RGB format, because the format is greatly influenced by changes in brightness. Therefore, we use the YCrCb color space. Y is the luma component. Cb and Cr are the blue-difference and red-difference chroma components. YCrCb format is robust to brightness changes. The color space was used because the human skin color has individual differences. The conversion from the RGB format to YCrCb format is performed using the following equations.

$$\begin{aligned}
 Y &= 0.229 \times R + 0.587 \times G + 0.114 \times B \\
 Cr &= 0.500 \times R - 0.419 \times G - 0.081 \times B \\
 Cb &= -0.169 \times R + 0.332 \times G + 0.500 \times B
 \end{aligned}$$



Fig.8 Input Image



Fig.9 Extraction

Fig.10 Labeling

The result of extracting skin color regions from Fig.8 are shown in Fig.9. Moreover, to remove the noise, the dilation and erosion processing is done. As shown Fig.9, it is understood that in addition to the hand, the face and a part of background are also extracted. Therefore, labeling is performed on the extracted area. The area of each area is calculated at the same time. The distance between the person performing the gestures and the robot is between 1 and 2 meters. Therefore, very large areas, for example the face, can be deleted because at this distance, they are too large to be hands. Moreover, obviously small areas (less than 40pixels) are likewise removed as noise compared with the hand. The result is shown in Fig.10.

3.2.2 Optical Flow

We now explain the technique for detecting the moving regions. The features are first detected from the input image, and are pursued by using optical flow. The Harris filter is used to capture the features. The Harris filter is one of the techniques for detecting edges, an effective characteristic for the tracking. The Lukas-Kanade algorithm was used for the tracking of the feature. This algorithm pursues a feature around the image. High-speed tracking is effective for the robotic operation if possible and real time processing is necessary. When the movement of the feature is detected, the image the result and the labeling result are compared, and the features region is assigned and recorded.

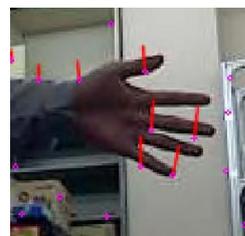


Fig.11 Optical Flow

3.3 Estimation of Hand's Direction

3.3.1 Feature Point Extraction

In this section we introduce how to estimate the direction of the hand. We assume that the estimated hand area to have a motion which we extract using optical flow. However, the hand area is small as shown in Fig.12. In the example of Fig.12, image size is 320x240[pixels] and hand area is about 50x40[pixels]. It is difficult to extract feature points. Therefore, we expand the area that contains the hand, to detect feature points thereafter.



Fig.12 Hand

We use the Harris operator as feature point's extractor. It can detect fingertips, because, it can detect corners. Hand area is enlarged to 300x300pixels to get more feature points. Fig.13 shows the result of extracted Harris feature points.

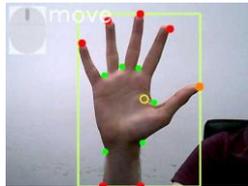
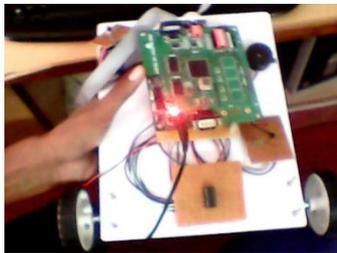


Fig.13 Feature Points

IV. Results and Discussion

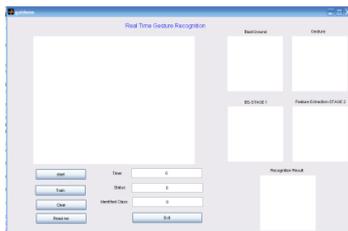
We conducted experiments to verify the effectiveness of our method.



These is the overview of the of the finger robo frontview

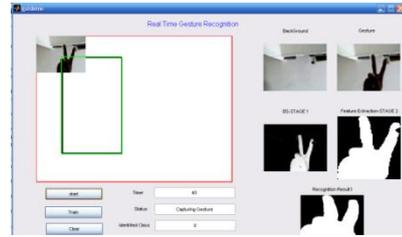
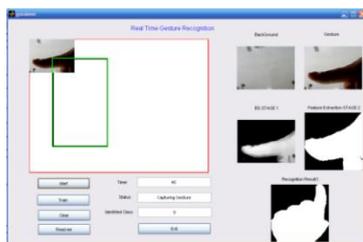


The fpga spartan3e kit which runs according to the command



The gesture recognition window where the feature points are extracted and functions accordingly.

The below windows describes the figure extraction which indicates the movement of the robot



V. Conclusions

In this paper, we proposed a simple operation method for a mobile robot using gesture recognition. We recognized hand gestures divided into motion and hand direction. We used optical flow to estimate the motion of the hand and distribution of feature points to estimate the hand direction. This method can be used to operate the robot in real time. However we encountered several problems. The first problem is color extraction. It is difficult to detect hand area when there is overlapping occur with the skin area in the background and the hand. The second problem is the feature points. In several images, we could not detect the feature points between the fingertip and fingers. In future, we must find solution to these problems and increase the types of gestures to enable complete robot control in a real environment.

References

- [1] Toshiya YAMADA, Kazunori UMEDA: Improvement of the Method of Operating a Mobile Robot by Gesture Recognition, <http://www.mech.chuo-u.ac.jp/umedalab/publication/s/pdf/intro/yamada.pdf>
- [2] Yuuji ASADA, Isao NISHIHARA, Shizuo NAKANO: A study of recognizing movement in hand gesture interface using image processing, ITE Technical Report Vol.34 No.18 (2010)
- [3] Yohei SATO, Kokichi SUGIHARA: An approach to real-time gesture recognition for human-PC interaction Guide to the Technical Report and Template, TECHNICAL REPORT OF IEICE NLC2004-112 (2005).
- [4] Kenichi FUKAYA : Manipulating a Mobile Robot by Hand Gestures, Hokkai-Gakuen University Fig.15 categorize 2.

FPGA Implementation of a New Parallel FIR Filter Structures

D. Suresh, K. Ramadevi, K. Raguram

*D.suresh (E.C.E), Associate professor, Associate professor, Pragati engineering college
 Surampalem, East Godavari Dist ,A.P.*

ABSTRACT : In recent days filters with large lengths are started to use. So parallel processing is essential at any cost. In this paper proposes new parallel FIR filter structures, which are beneficial to symmetric coefficients in terms of the hardware cost, under the condition that the number of taps is a multiple of 2 or 3. The proposed parallel FIR structures use symmetric property to reducing half the number of multipliers in sub filter section at the expense of additional adders in preprocessing and post processing blocks. Exchanging multipliers with adders is advantageous because adders weigh less than multipliers in terms of silicon area; in addition, the overhead from the additional adders in preprocessing and post processing blocks stay fixed and do not increase along with the length of the FIR filter, whereas the number of reduced multipliers increases along with the length of the FIR filter. Parallel FIR filter is essential, especially when the length of the filter is large.

Key words: Parallel FIR, preprocessing

I. INTRODUCTION

Finite impulse response (FIR) filters are the most popular type of filters implemented in software. This introduction will help you understand them both on a theoretical and a practical level. Filters are signal conditioners. Each functions by accepting an input signal, blocking pre-specified frequency components, and passing the original signal minus those components to the output. In a typical digital filtering application, software running on a digital signal processor (DSP) reads input samples from an A/D converter, performs the mathematical manipulations dictated by theory for the required filter type, and outputs the result via a D/A converter.

Some applications need the FIR filter to operate at high frequencies such as video processing, whereas some other applications request high throughput with a low-power circuit such as multiple-input multiple-output (MIMO) systems used in cellular wireless communication. Furthermore, when narrow transition-band characteristics are required, the much higher order in the FIR filter is unavoidable. For example, a 576-tap digital filter is used in a video ghost canceller for broadcast television, which reduces the effect of multipath signal echoes.

II. Finite Impulse Response

Filters can be classified in several different groups, depending on what criteria are used for classification. The two major types of digital filters are finite impulse response digital filters (FIR filters) and infinite impulse response digital filters (IIR).

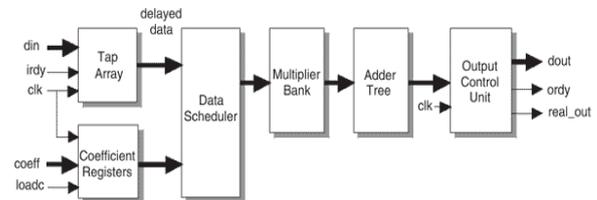


Figure.1 Digital filtering

Both types have some advantages and disadvantages that should be carefully considered when designing a filter. Besides, it is necessary to take into account all fundamental characteristics of a signal to be filtered as these are very important when deciding which filter to use. In most cases, it is only one characteristic that really matters and it is whether it is necessary that filter has linear phase characteristic or not.

Speech signal, for example, can be processed in the systems with non-linear phase characteristic. The phase characteristic of a speech signal is not of the essence and as such can be neglected, which results in the possibility to use much wider range of systems for its processing.

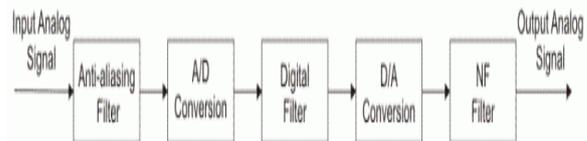


Figure 2. Digital filtering

The process of selecting the filter's length and coefficients is called filter design. The goal is to set those parameters such that certain desired stop band and pass band parameters will result from running the filter. Most engineers utilize a program such as MATLAB to do their filter design. But whatever tool is used, the results of the design effort should be the same:

A frequency response plot, like the one shown in Figure 1, which verifies that the filter meets the desired specifications, including ripple and transition bandwidth. The longer the filter (more taps), the more finely the response can be tuned. With the length, N , and coefficients, $float\ h[N] = \{ \dots \}$, decided upon, the implementation of the FIR filter is fairly straightforward. Listing 1 shows how it could be done in C. Running this code on a processor with a multiply-and-accumulate instruction (and a compiler that knows how to use it) is essential to achieving a large number of taps.

A. Ideal low-pass filter

FIR filters are digital filters with finite impulse response. They are also known as non-recursive digital filters as they do not have the feedback (a recursive part of a filter), even though recursive algorithms can be used for FIR filter realization

B. Window Method for FIR Filter Design

The window method for digital filter design is fast, convenient, and robust, but generally suboptimal. It is easily understood in terms of the convolution theorem for Fourier transforms, making it instructive to study after the Fourier theorems and windows for spectrum analysis.

$$[-N, N]$$

We would expect to be able to truncate it to the interval, for some sufficiently large N , and obtain a pretty good FIR filter which approximates the ideal filter. This would be an example of using the window method with the rectangular window. We saw in §4.3 that such a choice is optimal in the least-squares sense, but it designs relatively poor audio filters. Choosing other windows corresponds to tapering the ideal impulse response to zero instead of truncating it. Tapering better preserves the shape of the desired frequency response, as we will see. By choosing the window carefully, we can manage various trade-offs so as to maximize the filter-design quality in a given application. Window functions are always time limited. The window method always designs a finite-impulse-response (FIR) digital filter (as opposed to an infinite-impulse-response (IIR) digital filter). By the dual of the convolution theorem, point wise multiplication in the time domain corresponds to convolution in the frequency domain.

C. FIR And IIR Digital Filter Design

Based on combining ever increasing computer processing speed with higher sample rate processors, Digital Signal Processors (DSP's) continue to receive a great deal of attention in technical literature and new product design. The following section on digital filter design reflects the importance of understanding and utilizing this technology to provide precision stand alone digital or integrated analog/digital product solutions. By utilizing DSP's capable of sequencing and reproducing hundreds to thousands of discrete elements, design models can simulate large hardware structures at relatively low cost. DSP techniques can perform functions such as Fast-Fourier Transforms (FFT), delay equalization, programmable gain, modulation, encoding/decoding, and filtering.

- Filter weighting functions (coefficients) can be calculated on the fly, reducing memory requirements
- Algorithms can be dynamically modified as a function of signal input.

DSP represents a subset of signal-processing activities that utilize A/D converters to turn analog signals into streams of digital data. A stand-alone digital filter requires an A/D converter (with associated anti-alias filter), a DSP chip and a PROM or software driver. An extensive sequence of multiplication's and additions can then be performed on the digital data. In some applications, the designer may also want to place a D/A converter, accompanied by a reconstruction filter, on the output of the DSP to create an analog equivalent signal. A digital filter solution offering a 90 dB attenuation floor and a 20 kHz bandwidth can consist of up to 10 circuits occupying several square inches of circuit-board space and costing hundreds of dollars.

Digital filters process digitized or sampled signals. A digital filter computes a quantized time-domain representation of the convolution of the sampled input time function and a representation of the weighting function of the filter. They are realized by an extended sequence of multiplications and additions carried out at a uniformly spaced sample interval. Simply said, the digitized input signal is mathematically influenced by the DSP program. These signals are passed through structures that shift the clocked data into summers (adders), delay blocks and multipliers. These structures change the mathematical values in a predetermined way; the resulting data represents the filtered or transformed signal. It is important to note that distortion and noise can be introduced into digital filters simply by the conversion of analog signals into digital data, also by the digital filtering process itself and lastly by conversion of processed data back into analog.

When fixed-point processing is used, additional noise and distortion may be added during the filtering process because the filter consists of large numbers of multiplications and additions, which produce errors, creating truncation noise. Increasing the bit resolution beyond 16-bits will reduce this filter noise.

Instead of using a commercial DSP with software algorithms, a digital hardware filter can also be constructed from logic elements such as registers and gates, or an integrated hardware block such as an FPGA (Field Programmable Gate Array). Digital hardware filters are desirable for high bandwidth applications; the trade-offs are limited design flexibility and higher cost.

(1) Fixed-Point DSP and FIR (Finite Impulse Response) Implementations: Fixed-Point DSP processors account for a majority of the DSP applications because of their smaller size and lower cost. The Fixed-Point math requires programmers to pay significant attention to the number of coefficients utilized in each algorithm when multiplying and accumulating digital data to prevent distortion caused by register overflow and a decrease of the signal-to-noise ratio caused by truncation noise. The structure of these algorithms uses a repetitive delay-and-add format that can be represented as "DIRECT FORM-I STRUCTURE",

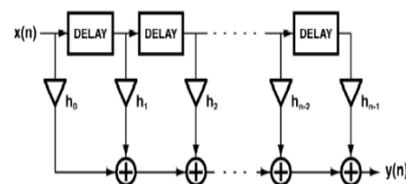


Figure 3 Transposed direct form FIR Filter

FIR (Finite Impulse Response) filters are implemented using a finite number "n" delay taps on a delay line and "n" computation coefficients to compute the algorithm (filter) function. The above structure is non-recursive, a repetitive delay-and-add format, and is most often used to produce FIR filters. This structure depends upon each sample of new and present value data. FIR filters can create transfer function that have no equivalent in linear circuit technology.

III. Window Technique:

The simplest technique is known as “Windowed” filters. This technique is based on designing a filter using well-known frequency domain transition functions called “windows”. The use of windows often involves a choice of the lesser of two evils. Some windows, such as the Rectangular, yield fast roll-off in the frequency domain, but have limited attenuation in the stop-band along with poor group delay characteristics. Other windows like the Blackman, have better stop-band attenuation and group delay, but have a wide transition-band (the band-width between the corner frequency and the frequency attenuation floor). Windowed filters are easy to use, are scalable (give the same results no matter what the corner frequency is) and can be computed on-the-fly by the DSP.

IV. The Equiripple Technique

An Equiripple or Remez Exchange (Parks-McClellan) design technique provides an alternative to windowing by allowing the designer to achieve the desired frequency response with the fewest number of coefficients. This is achieved by an iterative process of comparing a selected coefficient set to the actual frequency response specified until the solution is obtained that requires the fewest number of coefficients. Though the efficiency of this technique is obviously very desirable, there are some concerns.

- For equiripple algorithms some values may converge to a false result or not converge at all. Therefore, all coefficient sets must be pre-tested off-line for every corner frequency value.
- Application specific solutions (programs) that require signal tracking or dynamically changing performance parameters are typically better suited for windowing since convergence is not a concern with windowing.
- Equiripple designs are based on optimization theory and require an enormous amount of computation effort. With the availability of today’s desktop computers, the computational intensity requirement is not a problem, but combined with the possibility of convergence failure; equiripple filters typically cannot be designed on-the-fly within the DSP.

Analog filters beyond 10 poles are very difficult to realize and tend to be noisy

V. Digital to Analog Conversion (D/A)

As with input signals to A/D converters, waveforms created by D/A converters also exhibit errors. For each input digital data point, the D/A holds the corresponding value until the next sample period. Therefore, the output waveform exists as a sequence of steps. This output, a kind of “sample-and-hold” – is known as a “first-order hold.” In non-reconfigurable filters, these coefficients are constant and shift operation is done by hardwiring. The long tree of adders in multiplier implementation increases switching activity and physical capacitance and then power consumption.

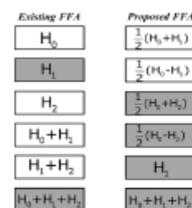


Fig.4 Implementation of coefficient

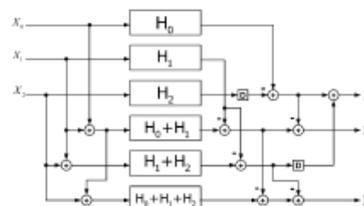


Fig.5 Parallel FIR filter architecture

VI. Proposed Reconfigurable Fir Filter Architecture

To utilize the symmetry of coefficients, the main idea behind the proposed structures is actually pretty intuitive, to manipulate the polyphase decomposition to earn as many subfilter blocks as possible which contain symmetric coefficients so that half the number of multiplications in the single subfilter block can be reused for the multiplications of whole taps, which is similar to the fact that a set of symmetric coefficients would only require half the filter length of multiplications in a single FIR filter. Therefore, for an N-tap 4-parallel FIR filter the total amount of saved multipliers would be the number of subfilter blocks that contain symmetric coefficients times half the number of multiplications in a single subfilter block decomposition to earn as many subfilter blocks as possible which contain symmetric coefficients so that half the number of multiplications in the single subfilter block can be reused for the multiplications of whole taps, which is similar to the fact that a set of symmetric coefficients would only require half the filter length of multiplications in a single FIR filter. Therefore, for an N-tap 3-parallel FIR filter the total amount of saved multipliers would be the number of subfilter blocks that contain symmetric coefficients times half the number of multiplications in a single subfilter block. As can be seen from the example above, two of three subfilter blocks from the proposed two-parallel FIR filter structure, H_0+H_1 and H_0-H_1 , are with symmetric coefficients now, as (8), which means the subfilter block can be realized by Fig. 4, with only half the amount of multipliers required. Each output of multipliers responds to two taps. Note that the transposed direct-form FIR filter is employed. Compared to the existing FFA two-parallel FIR filter structure, the proposed FFA structure leads to one more subfilter block which contains symmetric coefficients. However, it comes with the price of the increase of amount of adders in preprocessing and postprocessing blocks. In this case, two additional adders are required for $L=2$. Add/Sub control block. This block uses the sign bit of each sub-coefficient, and control the add/sub block. To implement the multiplication by zero for each subcoefficient, the multiplexer blocks are followed by AND gates, which is controlled by Mux control block. Three full add/sub blocks are used to combine the partial products of subcoefficients.

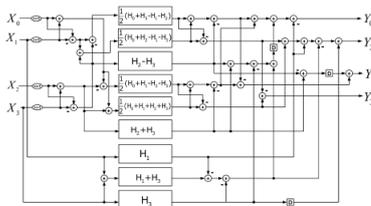


Figure. 6 Proposed parallel FIR filter architecture using four input.

VII. IMPLEMENTATION OF ALGORITHM

A primary objective of this project was to develop a synthesizable model for the AES128 encryption algorithm. Synthesis is the process of converting the register transfer level (RTL) representation of a design into an optimized gate-level netlist. This is a major step in ASIC design flow that takes an RTL model closer to a low-level hardware implementation.

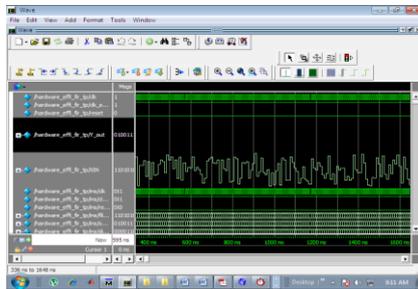


Figure7 .Simulated output.

A. Synthesis Timing Result

The synthesis tool optimizes the combinational paths in a design. In General, four types of combinational paths can exist in any design: [3]

- 1- Input port of the design under test to input of one internal flip-flop
- 2- Output of an internal flip-flop to input of another flip-flop
- 3- Output of an internal flip-flop to output port of the design under test
- 4- A combinational path connecting the input and output ports of the design under test

The last DC command in the script developed in previous section, instructs the tool to report the path with the worst timing. In this case, the path with the worst timing is a combinational path of type two. The delay associated with this path is the summation of delays of all combinational gates in the path plus the *Clock-To-Q* delay of the originating flip-flop, which was calculated as 24.09ns.

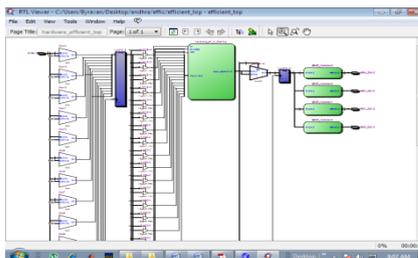


Figure 8.RTL Schematic report

By considering the setup time of the destination flip-flop in this path, which is 0.85ns, the

40MHz clock signal satisfies the worst combinational path delay. The delays of combinational gates, setup time of flip-flops and *Clock-To-Q* values are derived from the LSI_10k library file that was used for the mapping step during synthesis

B. Synthesis Area Result

The synthesis area report shows the total number of cells and nets in the netlist. It also uses the area parameter associated with each cell in the LSI_10K library file, to calculate the total combinational and sequential area of the netlist. The total area of the gate level netlist is unknown since it depends on total area of the interconnects, which itself is a function of the wiring load model used in physical design. The total cell area in the netlist is reported as 22978 units, which is the sum of combinational and sequential areas.

Flow Summary	
Flow Status	Successful - Sun May 20 08:53:24 2012
Quartus II Version	11.0 Build 208 07/03/2011 SP 1.53 Web Edition
Revision Name	efficient_top
Top-level Entity Name	hardware_efficient_top
Family	Cyclone III
Device	EP3C16F48K4C6
Timing Models	Final
Total logic elements	2,925 / 15,408 (19 %)
Total combinational functions	2,754 / 15,408 (18 %)
Dedicated logic registers	355 / 15,408 (2 %)
Total registers	355
Total pins	35 / 347 (10 %)
Total virtual pins	0
Total memory bits	0 / 516,096 (0 %)
Embedded Multiplier 9-bit elements	0 / 112 (0 %)
Total PLLs	0 / 4 (0 %)

Figure 9.Flow summary report

To enforce the synthesis tool to create the most compact netlist, the area of the gate level netlist was constrained to zero during the synthesis process. As a result, the only constraint violation, which is expected, is related to the area as shown bellow:

C. Performance Report

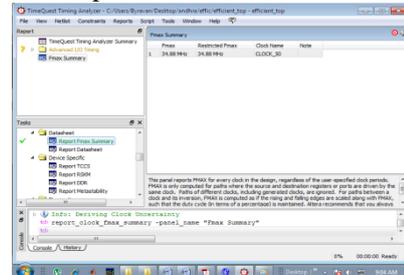


Figure10 .Fmax. summary report for slow corner.

VIII. CONCLUSION

The proposed new structure exploits the nature of even symmetric coefficients and save a significant amount of multipliers at the expense of additional adders. Since multipliers outweigh adders in hardware cost, it is profitable to exchange multipliers with adders. Moreover, the number of increased adders stays still when the length of FIR filter becomes large, whereas the number of reduced multipliers increases along with the length of FIR filter. Consequently, the larger the length of FIR filters is, the more the proposed structures can save from the existing FFA structures, with respect to the hardware cost. Overall this paper proved that for larger filter length area consumption of proposed filter is far better than any other existing method.

REFERENCES

- [1] D. A. Parker and K. K. Parhi, "Low-area/power parallel FIR digital filter implementations," *J. VLSI Signal Process. Syst.*, vol. 17, no. 1, pp. 75–92, 1997.
- [2] J. G. Chung and K. K. Parhi, "Frequency-spectrum-based low-area low-power parallel FIR filter design," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 9, pp. 444–453, 2002.
- [3] K. K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation*. New York: Wiley, 1999.
- [4] Z.-J. Mou and P. Duhamel, "Short-length FIR filters and their use in fast nonrecursive filtering," *IEEE Trans. Signal Process.*, vol. 39, no. 6, pp. 1322–1332, Jun. 1991.
- [5] J. I. Acha, "Computational structures for fast implementation of L-path and L-block digital filters," *IEEE Trans. Circuit Syst.*, vol. 36, no. 6, pp. 805–812, Jun. 1989.
- [6] C. Cheng and K. K. Parhi, "Hardware efficient fast parallel FIR filter structures based on iterated short convolution," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 8, pp. 1492–1500, Aug. 2004.
- [7] C. Cheng and K. K. Parhi, "Further complexity reduction of parallel FIR filters," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS 2005)*, Kobe, Japan, May 2005.
- [8] C. Cheng and K. K. Parhi, "Low-cost parallel FIR structures with N -stage parallelism," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 2, pp. 280–290, Feb. 2007.
- [9] I.-S. Lin and S. K. Mitra, "Overlapped block digital filtering," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 43, no. 8, pp. 586–596, Aug. 1996.
- [10] "Design Compiler User Guide," ver. B-2008.09, Synopsys Inc., Sep. 2008.

Integration of Histogram Insertion Measure Method for Achieve Effecient Retreival of Images

Mrs. V. Indrani¹, Mrs. K. Prathima², Mr. B. Vikas³, Mrs. Ch. Roopasree⁴
 (Dept of CSE, Vignan's Institute Management and Technology for Women, Hyderabad, Andhra Pradesh, India)

ABSTRACT : Retrieving information in the form of images had been wide spreading research now a days. Knowledge transmission is achieved through information retrieval in both the forms of Images and Texts, Presently researchers and mainly focusing on this. Mainly in our paper we are going to incorporate the method of Histogram Intersection measure to evaluate the database image to Query image, which helps to measure the over-all similarities between all the images in the Knowledge Database. Hence we prove that our approach is accurate by incorporating the all the local properties of the texture histogram of the images.

Keywords: Image Retrieval, Texture Features, Knowledge Transmission, Query image, Local Properties

I. INTRODUCTION

To date, image and video storage and retrieval systems have typically relied on human supplied textual annotations to enable indexing and searches. The text-based indexes for large image and video archives are time consuming to create. They necessitate that each image and video scene is analyzed manually by a domain expert so the contents can be described textually. The language-based descriptions, however, can never capture the visual content sufficiently. For example, a description of the overall semantic content of an image does not include an enumeration of all the objects and their characteristics, which may be of interest later. A content mismatch occurs when the information that the domain expert ascertains from an image differs from the information that the user is interested in. A content mismatch is catastrophic in the sense that little can be done to approximate or recover the omitted annotations. In addition, a language mismatch can occur when the user and the domain expert use different languages or phrases. Because text-based matching provides only hit-or-miss type searching, when the user does not specify the right keywords the desired images are unreachable without examining the entire collection.

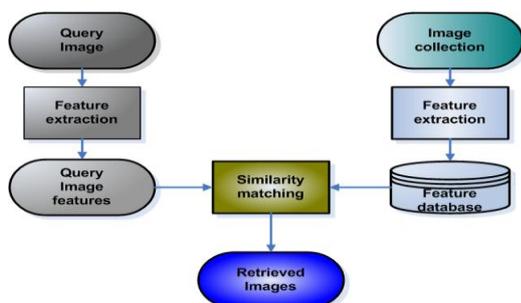


Fig.1: The Architecture of CBIR Technique

The prime requirement for Retrieval systems is to be able to display images relating to a named query image. The text indexing is often limited, tedious and subjective for

describing image content. So there is increasing interest in the use of CBIR techniques. The problems with text-based access to images have prompted increasing interest in the development of image based solutions. This is more often referred to as Content Based Image Retrieval (CBIR) as shown in Fig.1. Content Based Image Retrieval relies on the characterization of primitive features such as color, shape and texture that can be automatically extracted from the images themselves. Queries to CBIR system are most often expressed as visual exemplars of the type of the image or image attributed being sought. For Example user may submit a sketch, click on the texture pallet, or select a particular shape of interest. This system then identifies those stored images with a high degree of similarity to the requested feature.

Digital imaging has become the standard for all image acquisition devices. So there is an increasing need for data storage and retrieval. With lakhs of images added to the image database, not many images are annotated with proper description. So many relevant images go unmatched. The most widely accepted content-based image retrieval techniques cannot address the problems with all images, which are highly specialized .Our approach Histogram based Image Retrieval using Texture Feature retrieves the relevant images based on the texture property. We also provide an interface where the user can give a query image as an input. The texture feature is automatically extracted from the query image and is compared to the images in the database retrieving the matching images.

II. BACKGROUND & PREVIOUS WORK

The goal of Content-Based Image Retrieval (CBIR) systems is to operate on collections of images and, in response to visual queries, extract relevant image. The application potential of CBIR for fast and effective image retrieval is enormous, expanding the use of computer technology to a management tool.

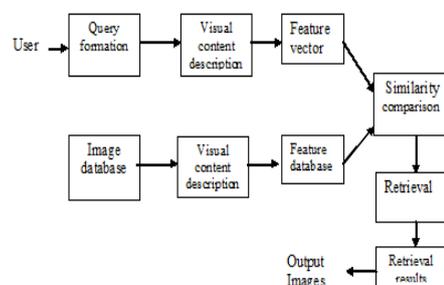


Fig.2: Procedure for content based image retrieval system

CBIR operates on the principle of retrieving stored images from a collection by comparing features automatically extracted from the images themselves. The commonest features used are mathematical measures of color, texture or shape. A typical system allows users to formulate queries

by submitting an example of the type of image being sought, though some offer alternatives such as selection from a palette or sketch input. The system then identifies those stored images whose feature values match those of the query most closely, and displays thumbnails of these images on the screen.

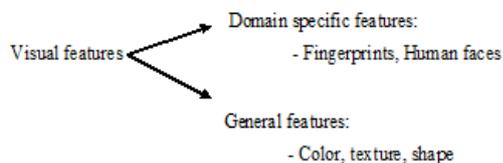


Fig.3: Approaches of CBIR

Some of the most commonly used types of features used for image retrieval are as follows:

a) Colour retrieval

Several methods for retrieving images on the basis of color similarity have been described in the literature, but most are variations on the same basic idea. Each image added to the collection is analyzed to compute a color histogram, which shows the proportion of pixels of each color within the image. The color histogram for each image is then stored in the database. At each time, the user can either specify the desired proportion of each color (75% olive green and 25% red, for example), or submit an example image from which a color histogram is calculated. Either way, the matching process then retrieves those, which a color histogram is calculated. Either way, the matching process then retrieves those images whose color histograms match those of the query most closely.

b) Texture retrieval

The ability to match on texture similarity can often be useful in distinguishing between areas of images with similar color (such as blue sky and sea or green leaves and grass). A variety of techniques has been used for measuring texture similarity; the best established rely on comparing values of what are known as second-order statistics calculated from query and stored images. Essentially, these calculate the relative brightness of selected pairs of pixels from each image. From these it is possible to calculate measures of image texture such as the degree of contrast, coarseness, directionality and regularity or periodically, directionality and randomness.

Texture queries can be formulated in a similar manner to color queries, by selecting examples of desired texture a palette, or by supplying an example query image. The system then retrieves images with texture measures most similar in value to the query.

c) Shape retrieval

Two major steps are involved in shape feature extraction. They are object segmentation and shape representation.

Object segmentation: Segmentation is very important to Image Retrieval. Both the shape feature and the layout feature depend on good segmentation allow fast and efficient searching for information of a user's need.

d) Shape representation

In image retrieval, depending on the applications, some requires the shape representation to be invariant to translation, rotation, and scaling. In general, the shape representations can be divided into two categories, boundary-based and region-based. The former uses only the outer boundary of the shape while the latter uses the entire shape region.

III. CONCEPT OF TEXTURE FEATURE

Texture is one of the crucial primitives in human vision and texture features have been used to identify contents of images. Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. One crucial distinction between color and texture features is that color is a point, or pixel, property, whereas texture is a local-neighbourhood property. As a result, it does not make sense to discuss the texture content at pixel level without considering the neighbourhood.

The texture is a property inherent to the surface. Various parameters or textural characteristics describe it. They are:

- The Granularity which can be rough or fine
- The Evenness which can be more or less good
- The Linearity
- The directivity
- The repetitiveness
- The contrast
- The order
- The connectivity

The other characteristics like color, size, and shape also must be considered. The Methodologies used for analysis of the texture are as follows

i. Texture spectrum method

The basic concept of texture spectrum method was introduced by H1 and Wang. The texture can be extracted from the neighborhood of 3 X 3 window which constitute the smallest unit called 'texture unit'. The neighborhood of 3 X 3 consists of nine elements respectively as $V = \{ V_1, V_2, V_3, V_4, V_0, V_5, V_6, V_7, V_8 \}$ where V_0 is the central pixel value and $V_1 \dots V_8$ are the values of neighboring pixels within the window. The corresponding texture unit for this window is then a set containing eight elements surrounding the central pixel, represented as:

$$TU = \{ E_1, E_2, E_3, E_4, E_0, E_5, E_6, E_7, E_8 \}$$

Where E_i is defined as: $E_i = 0$ if $V_i < V_0$

$$1 \text{ if } V_i = V_0$$

$$2 \text{ if } V_i > V_0$$

And the element E_1 occupies the corresponding V_1 pixel. Since each of the eight element of the texture units has any one of three values (0, 1, or 2)

$$NTU = \sum E_i * 3^{(i-1)} \quad [\text{For } i=1 \text{ to } 8]$$

Where NTU is the texture unit value. The occurrence distribution of texture unit is called the texture spectrum (TS). Each unit represents the local texture information of 3X3 pixels, and hence statistics of all the texture units in an image represent the complete texture aspect of entire image.

ii. Cross diagonal texture spectrum

AL-Jan obi (2001) has proposed a cross-diagonal texture matrix technique. In this method the eight neighboring pixels of 3 X 3 widows is broken up into two groups of four elements each at cross and diagonal positions. These groups are named as Cross Texture Unit (CTU) and Diagonal Texture Unit (DTU) respectively. Each of the four elements of these units is assigned a value (0, 1, and 2) depending on the gray level difference of the corresponding pixel with that of the central pixel of 3X3 window. These texture units have values from 0 to 80 (34, i.e 81 possible values).

Cross Texture Unit (CTU) and Diagonal Texture Unit (DTU) can be defined as:

$$NCTU = \sum E_{ci} * 3 (i - 1) \quad [\text{For } i=1 \text{ to } 4]$$

$$NDTU = \sum E_{di} * 3 (i - 1) \quad [\text{For } i=1 \text{ to } 4]$$

Where NCTU and NDTU are the cross texture and diagonal texture unit values respectively; E_{ci} and E_{di} are the i th elements of texture unit

The texture unit (CTU orDTU) value can range between: (0-240)

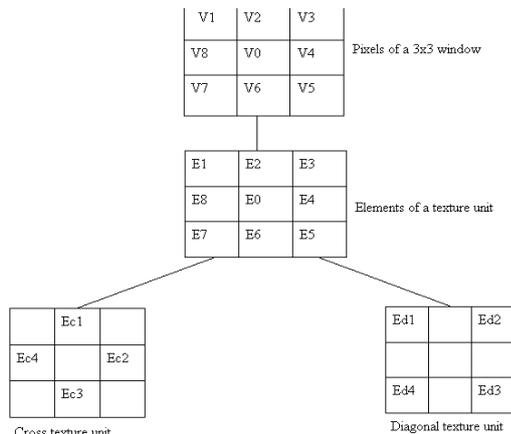


Fig.4: Formation of cross-diagonal texture units

i. Modified texture spectrum

In the proposed method, Nctu and Ndtu values have been evaluated which range from 0 to 80. For each type of texture unit, there can be four possible ways of ordering, which give four different values of CTU and DTU.

$$NTU = NCTU * NDTU$$

$$Ntu = Nctu+Ndtu$$

$$Ntu = Nctu-Ndtu$$

Where Nctu and Ndtu are the ordering ways for evaluation of Nctu and Ndtu.

After obtaining the CDTM values of 3*3 windows through entire image the occurrence frequency of each CDTM values are recorded. For the texture units having same CDTM values, two different procedures have been carried out to replace the pixel values of these units. The texture unit value can range between :(0-480).

ii. Texture spectrum with threshold

The texture spectrum method with threshold is intended to make difference between the values of neighborhood matrix which are very close to the central pixel value and those the rest. In this method the texture unit matrix is represented as:

$$TU = \{ E_1, E_2, E_3, E_4, E_0, E_5, E_6, E_7, E_8 \}$$

Where E_i is defined as: $E_i = 0$ if $V_i \leq (V_0 + t)$

$$1 \text{ if } V_i > (V_0 + t)$$

Where t is the threshold value.

$$NTU = \sum E_i * 2 (i - 1) \quad [\text{For } i=1 \text{ to } 8]$$

The texture unit value can range between (0-254).

iii. Reduced texture unit

In this method the range of texture unit values are (0,1). As the range is decreased the memory required to compute texture unit value also reduces. In this method

$$TU = \{ E_1, E_2, E_3, E_4, E_0, E_5, E_6, E_7, E_8 \}$$

Where E_i is defined as: $E_i = 0$ if $V_i \leq V_0$

$$1 \text{ if } V_i > V_0$$

Where t is the threshold value.

$$NRTU = \sum E_i * 2 (i - 1) \quad [\text{For } i=1 \text{ to } 8]$$

The texture unit value can range between (0-254).

iv. Splitting texture unit matrix into rows and columns

In this approach the texture unit matrix is split into 3 separate rows/columns. Texture unit value is calculated separately for each row/column. Later all the 3 texture unit values are added to get a single texture unit value. By doing this the texture unit value can be limited to 42. Thus memory and computation time can be saved.

E_1	E_2	E_3
E_8	E_0	E_4
E_7	E_6	E_5

Splitting into columns:

E_{11}	E_{21}	E_{31}
E_{12}	E_{22}	E_{32}
E_{13}	E_{23}	E_{33}

Here $[E_{11}- E_{13}]$ are the first column values of the texture unit matrix denoted as TU_1 . Similarly $[E_{21}- E_{23}]$ and $[E_{31}- E_{33}]$ denote second (TU_2) and third (TU_3) columns of texture unit matrix respectively.

The texture unit value is calculated separately for each texture unit matrix (j) as:

$$N TU_j = \sum E_{ji} * 2 (i - 1) \quad [\text{For } i=1 \text{ to } 3]$$

The final texture unit value is evaluated as:

$$N TU = \sum N TU_j \quad [\text{For } j=1 \text{ to } 3]$$

The texture unit value can range between (0-42).

Splitting into rows:

E ₁₁	E ₁₂	E ₁₃
E ₂₁	E ₂₂	E ₂₃
E ₃₁	E ₃₂	E ₃₃

Here [E11- E13] are the first row values of the texture unit matrix denoted as TU1. Similarly [E21- E23] and [E31- E33] denote second (TU2) and third (TU3) rows of texture unit matrix respectively.

The texture unit value is calculated separately for each texture unit matrix (j) as:

$$N TU_j = \sum E_{ji} * 2 (i - 1) \quad [\text{For } i=1 \text{ to } 3]$$

The final texture unit value is evaluated as:

$$N TU = \sum N TU_j \quad [\text{For } j=1 \text{ to } 3]$$

The texture unit value can range between (0-42).

IV. HISTOGRAM INTERSECTION APPROACH

To overcome the disadvantages of Euclidean distance we taken histogram intersection measure. The histogram intersection was investigated for color image retrieval by swain and Ballard. Their objective was to find known objects within images using color histograms. When the object (q) size is less than the image (t) size, and the histograms are not normalized, then |hq| <= |ht|. The intersection of histograms hq and ht is given by:

$$d_{q,t} = 1 - \frac{\sum_{m=0}^{M-1} \min(h_q[m], h_t[m])}{|h_q|}$$

Where |h| = Σ h[m] [for m=0 to M-1]. The above equation is not a valid distance metric since it is not symmetric hq,t not equal to dt,q. However that equation can be modified to produce a true distance metric by making it symmetric in hq and ht as follows:

$$d^1_{q,t} = 1 - \frac{\sum_{m=0}^{M-1} \min(h_q[m], h_t[m])}{\min(|h_q|, |h_t|)}$$

Alternatively when the histograms are normalized such that |hq|=|ht|, both equations are true distance metrics. When |hq|=|ht| that D1(q,t)=dq,t and the Histogram Intersection is given by

$$D1(q, t) = \sum |h_q[m] - h_t[m]| \quad [\text{For } m=0 \text{ to } M-1]$$

V. DESIGN & ANALYSIS OF THE EXPERIMENT

Class Diagram models class structure and contents using design elements such as classes, packages and objects as shown in Fig.5. It also displays relationships such as containment, inheritance, associations and others.

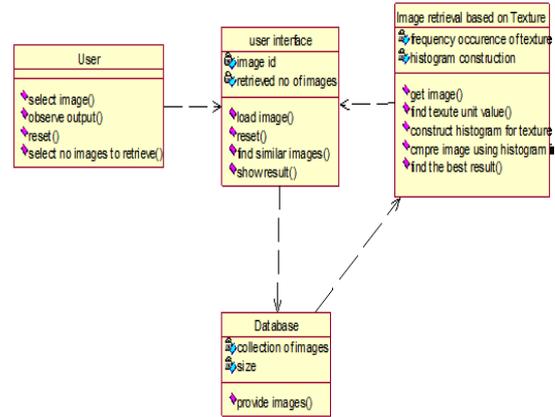


Fig.5: Interpretational Class Diagram for User interaction

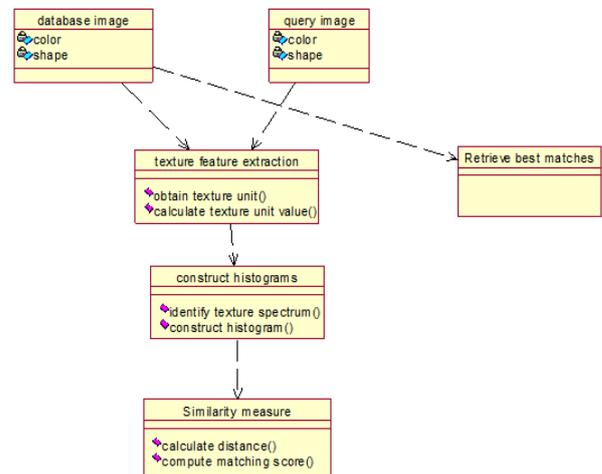


Fig.6: Interpretational Class Diagram for Image Retrieval

Sequence Diagram displays the time sequence of the objects participating in the interaction as shown in Fig.6. This consists of the vertical dimension (time) and horizontal dimension (different objects).

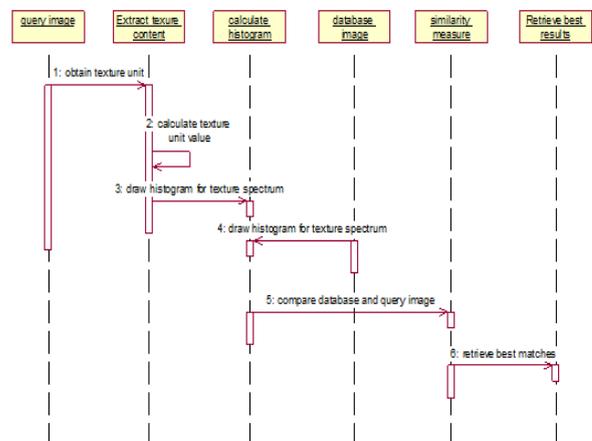


Fig.7: Interpretational Sequence Diagram of Database search

VI. RESULTS

The below are the results obtained from the experiment

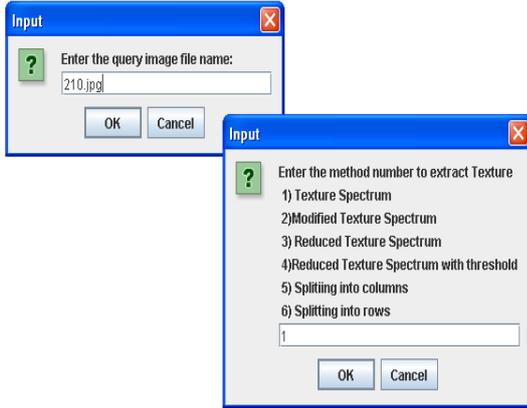


Fig.8: Texture Spectrum Method

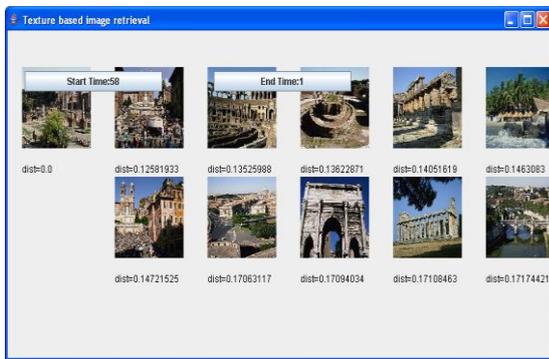


Fig.9: Modified Texture Spectrum Method

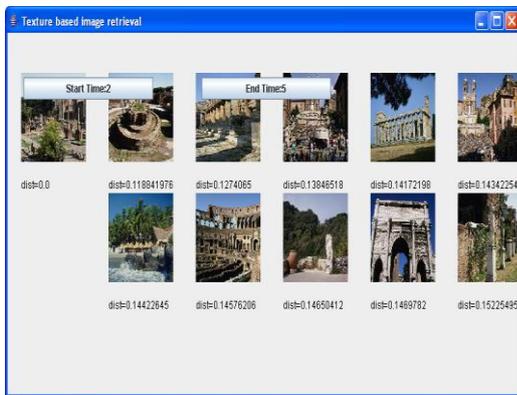


Fig.10: Reduced Texture Spectrum Method

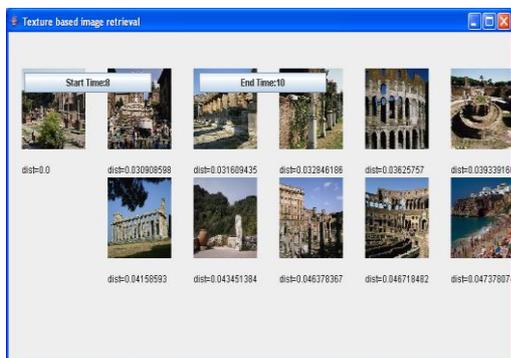


Fig.11: Reduced Texture Spectrum with Threshold

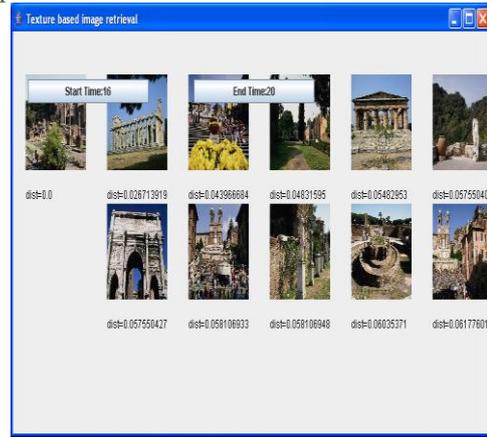


Fig.12: Splitting into Rows

VII. CONCLUSION

We presented a new idea of achieving Knowledge transmission of image retrieval through Histogram based Image Retrieval using Texture Feature system with different methods of extracting texture feature in both the forms of Images and Texts .We mainly concentrated and successes to incorporate the method of Histogram Intersection measure to evaluate the database image to Query image, which helped to calculate over all similarities between all the images in the Database.

REFERENCES

- [1] A. K. Jain and R. Dubes. Algorithms for clustering data. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988
- [2] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. IEEE Transactions Systems on Man and Cybernetics, 3(6):610-621, 1973.
- [3] W. Y. Ma and B. S. Manjunath. Netra: a toolbox for navigating large image databases. In International Conference on Image Processing (ICIP), pages 568–571, Oct. 1997.
- [4] T. Lehmann, B. Wein, J. Dahmen, J. Bredno, F. Vogelsang, and M. Kohlen. Content-Based Image Retrieval in Medical Applications: A Novel Multi-Step Approach. In International Society for Optical Engineering (SPIE), volume 3972(32), pages 312–320, Feb. 2000.
- [5] A. Bovik, M. Clark, and W. Geisler. Multichannel texture analysis using localized spatial filters. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12:55–73, 1990. T. S. Huang, S. Mehrotra, and K. Ramchandran.
- [6] F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Protopapas. Fast and effective retrieval of medical tumor shapes. IEEE Transactions on Knowledge and Data Engineering, 10(6):889–904, 1998.
- [7] A. Kak and C. Pavlopoulou. Content-Based Image Retrieval from Large Medical Databases. In 3D Data Processing, Visualization, Transmission, Padova, Italy, June 2002.

- [8] W. Chu, C. Hsu, A. Cardenas, and R. T aira. Knowledge-based image retrieval with spatial and temporal constructs. *IEEE Transactions on Knowledge and Data Engineering*, 10(6):872–888, 1998.
- [9] D. Clausi and M. Jernigan. Designing Gabor filters for optimal texture separability. *Pattern Recognition*, 33:1835–1849, Jan. 2000.
- [10] J. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, and R. Jain. Virage image search engine: An open framework for imagemanagement. In *SPIE Storage and Retrieval for Image and Video Databases*, volume 2670, pages 76–87, San Diego/La Jolla, CA, Jan. 1996.
- [11] D. Comaniciu, P. Meer, D. Foran, and E. Medl. Bimodal system for interactive indexing and retrieval of pathology images. In *Workshop on Applications of Computer Vision*, pages 76–81, Princeton, NJ, Oct. 1998.
- [12] Multimedia Analysis and Retrieval System (MARS) project. In *Proceedings of 33rd Annual Clinic on Library Application of Data processing - Digital Image Access and Retrieval*, pages 100–117, Urbana/Champaign, Illinois, 1996

Author List



Mrs. V. Indrani received M.Tech (CST with Bioinformatics) from Andhra University, Visakhapatnam. Presently she is working as Associate Professor at Vignan Institute Management and Technology for Women, Hyderabad, Andhra Pradesh, India. She is having 7+ years of teaching experience in the field of Computer Science and Engineering.



Mrs. K. Prathima received M.Tech (CSE) from Osmania University, Hyderabad. Presently she is working as Assistant Professor at Vignan's Institute of Management and Technology for Women, Hyderabad, Andhra Pradesh, India. She is having 7+ years of teaching experience in the field of Computer Science and Engineering.



Mr. B. Vikas received M.Tech (Bioinformatics) from JNT University, Hyderabad. Presently he is working as Assistant Professor at Vignan's Institute of Management and Technology for Women, Hyderabad, Andhra Pradesh, India. He is having 2+ years of teaching experience in the field of Computer Science and Engineering.



Mrs. Ch. Roopasree received M.Tech (Bioinformatics) from JNT University, Hyderabad. Presently she is working as Assistant Professor at Vignan's Institute of Management and Technology for Women, Hyderabad, Andhra Pradesh, India. She is having 3+ years of teaching experience in the field of Computer Science and Engineering.

Design and Implementation of A TALKING LCD Display

Rohit Kumar¹, B. Malla Reddy², Prof. N. Bhoopal³
^{1,2}Electronics and Communication Engineering Department, Mizoram University
³Electrical and Electronics Engineering Department, BVRIT

Abstract: the most electronic device conveys the information by Liquid Crystal Display (LCD). Visually Impaired people cannot use those type of devices. We present a design and construction of a Talking LCD display Clock. The design incorporates audio and visual function. The reproduction of human voice accomplished by pre-recording on speech module, a simple circuit of control unit makes the design affordable, practice, economical and user friendly.

Keywords: speech module (APR-9600), delay time, RTC (DS1307), LCD display, storage time, Proteus Isis, Visually impaired

I. INTRODUCTION

This project is completely independent. Unlike the existing solution [3], it depends on extremely powerful automatic text recognition and processing that can process text and transform it into speech or an alarm. The second aim of the project is to develop an alarm/speech display for visually impaired. The application facilities the impaired persons to customize a reminder alarm within it by customized user voice recording system. For example, if a person with visual disability wants to set an alarm or a reminder, he will not able to communicate and set. Therefore it is significant for visually impaired to record and receive information at all times. Many devices are emerging to address the related problem of reading text in printed documents [2, 16], but they are not designed to tackle the challenge of finding and reading characters in appliance displays [1, 14, 15]. Many devices have been built for security alarm [9, 10, 11] and communication aid [12, 13] visually impaired persons. We designed a talking LCD display clock which is can be easily operated by the visually impaired people. For performing any operation like to know the time one has to press the button, to set the alarm one has to go through the instructions spoken by the system. For conforming any option mentioned in the instruction one has to just press the button. The block diagram of the design is shown below:

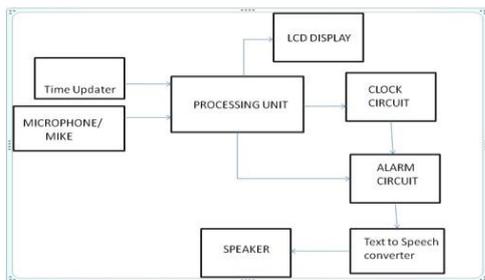


Fig 1.1: Block diagram of Talking LCD display Clock for Visually Impaired

II. DESGIN OVERVIEW

1. Flow Chart of Control in the System

The designed system will work as the flow chart shown in the fig 2.1.1. The time and date for the clock is received from real time clock (RTC-DS1307) and then transferred to the LCD display. The watch dog timer occur in the design when more than one button is pressed before the execution of command of first button is completed. In the case, the execution of command of second button will depend upon the executing address of APR9600. The system will not execute the second command till the execution of that address is completed.

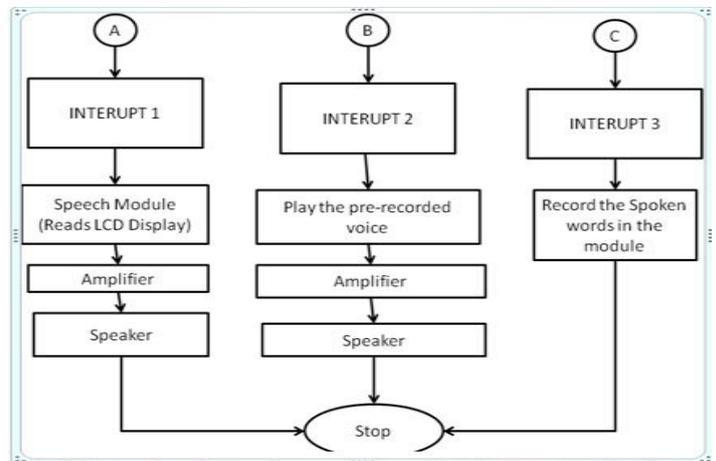
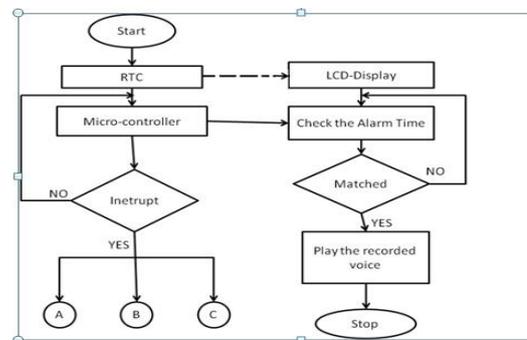


Fig 2.1.1: Flowchart for operation of Taking LCD display clock

2. Hardware Set Up

As shown in fig.2.2.1 the signal of time and date is received from RTC (DS1307) in micro-controller PIC18F4550 and then bypassed to LCD which is parallel connected to the micro-controller. The LCD display besides its seven pin for display, has three more pins namely RS for reset, RW for read-write operation and E for enable purpose. The speech module APR9600 has pre-recorded human voice and one address free for the recording of the remainder tone. The 8 address pins of the speech module are connected with the 8 pins of micro- controller. The keypad connected to the micro-controller has only first row active buttons.

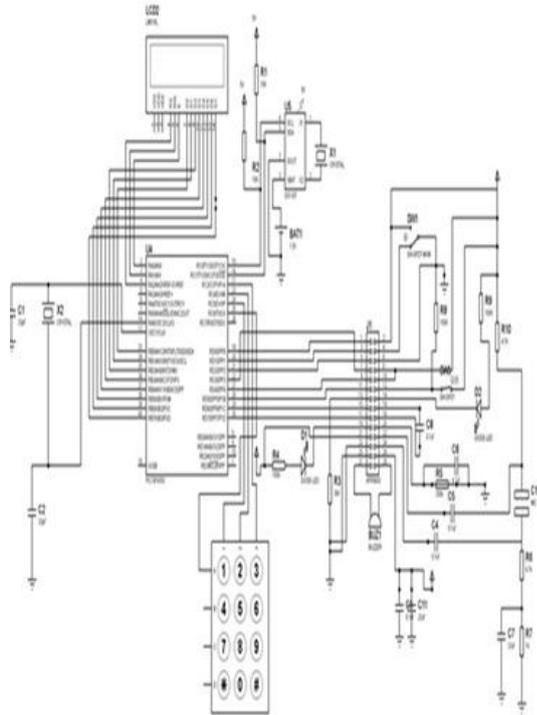


Fig 2.2.1: Schematics of Talking LCD display clock

The first button indicated as 1 is used for the purpose of listening time and date, the second button indicated as 2 is used for setting the alarm time and the third button indicated as 3 is used for recording the reminder for alarm.

III. MODELLING

1. RECORDED VOICE

The pre-recorded voice is stored at the eight memory locations in APR9600 speech module as shown in fig 3.1.1 below:

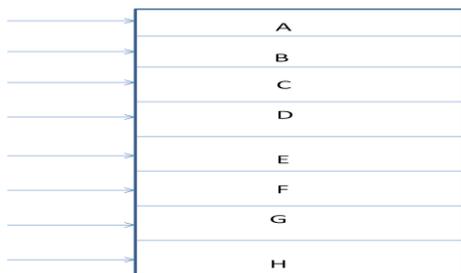


Fig 3.1.1: Memory location in APR9600

The pre-recorded human voice can be retrieved from the speech module by making the addressing pin of the memory location high[20]. Suppose we want the sound stored at memory location A as shown in fig 2.C.1.1 to be retrieved then the pin addressing memory location A is made high.

The sound from the module is not received continuously if we want to get from every location at the instant, as shown in fig 3.1.2 below:

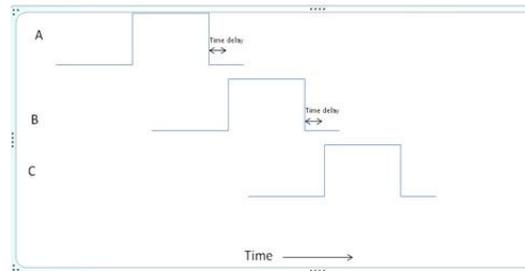


Fig 3.1.2: Delay time in operation of two consecutive memory locations.

Although the delay time is very small but plays important role in case of two interrupt occur due to consecutive pressing of buttons before the full command of first interrupt is executed.

The storage time in the module depend upon the frequency we want to use as shown in equation 2.1 below as:

$$Storage\ time = \frac{k}{f} \dots\dots\dots (2.1)$$

Here, 'k' is the proportionality constant having value of 2×10^{-3} at room temperature and depends upon the temperature and [5, 6], 'f' is the frequency band at which we want to record the voice; it is a function of frequency of oscillator.

2. REAL TIME CLOCK

RTC (DS1307) is a timer and has a life of about hundred years and has consumption of about 500nA [17]. RTC is connected with the micro-controller by its SDA and SCL port [7, 18]. The timing diagram of both the port is shown in fig 3.2.1. The synchronisation of DS1307 with the system is very efficient [8].

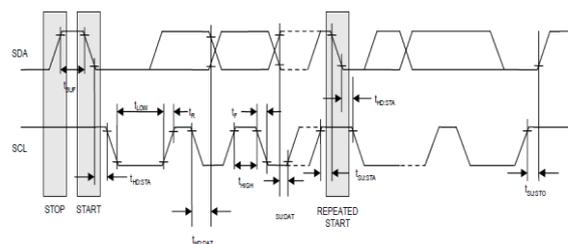


Fig 3.2.1: Timing diagram in RTC

3. POWER SUPPLY UNIT

The 5V/ 2Amp adapter is used to provide the constant power of 5V to the system at 133kHz frequency instead of regulated power supply as there is less power loss in the power supply unit.

IV. RESULTS AND DISCUSSION

1. Simulation Result

By switching ON the power supply of the device the LCD display starts displaying the time and date on the screen. The simulation software used is Proteus Isis and the programming is done on software Mikro C Pro for Pic and then dumped in the micro-controller.

2. Result of Implementation of Hardware

The alarm system work properly for one alarm and can be incremented if desired. The condition of hanging of system is avoided in the design with help of APR9600.

The complexity of design and programme increases with the increase in number of APR9600 used.

Same technique can be used for the purpose of security alarm and temperature measurement can be included in the same manner.

3. Conclusion

The system can be used very well within a large group of visually impaired person. The algorithm can be implemented in any kind of talking display. Voice activating feature can make it very useful for paralysed persons.

ACKNOWLEDGMENT

The authors pay their sincere gratitude to the management of BVRIT for funding this technical work. We also are grateful to ATL for fulfilling the technical requirement of the project. Rohit Kumar and B.Malla Reddy extends their heartily thanks to Hardware Engineer, ATL-BVRIT for his valuable time and effort for scrutinizing the design of the project. RK and BMR is obliged to LLK Singh , NP Maithy, Reshmi Maity and Abhijoti Gosh from Mizoram University to help us as when required.

REFERENCES

- [1] Paul Blenkhorn .,"Designing products that speak:lessons from talking systems for blind people ,computing and control engineering journal,2009
- [2] Xuanzhang , Cesar Ortega-Sanchez , Iain murray., IEEE ,2009,978-0-7695-3474-9/08
- [3] Marius A.Beukes, , Beatrys M. lacquet., IEEE AFRICON ,2004 ,0-7083-8605-1-2004
- [4] GPS Based Voice Alert System For Blind – Rishabb Gulati, IJSER, volume-2, Issue-1, Jan-2011
- [5] Remote Operation of Household Gadget using mobile phone- Debapratim Sakar, Saikat Das, Dept. ECE, Technoindi College
- [6] Frequency Hopping Synchronization Scheme Based on Real-Time Clock - Novi Sad, Serbia , April 11-April 13, ISBN: 978-0-7695-4664-3
- [7] A Multifunction domestic Alert System for the Deaf and blind -Robert I. Damper, *Senior Member, IEEE*, and Mike D. Evans, VOL. 3, NO. 4, DECEMBER 1995
- [8] J. Hurvitz and R. Carmen, *Special Devices for Hard of Hearing, Deaf and Deaf-Blind Persons*. Boston, MA: Little, Brown and Co., 1981.
- [9] T. A. Dumas, "Aids for the deaf," in *Electronic Devices for Rehabilitation*, J. G. Webster, A. M. Cook, W. J. Tompkins, G. C. Vanderheiden, Eds. London, England Chapman and Hall, 1985, pp. 148-174.
- [10] Portable communication aid for deaf-blind people- Mu-Chun Su, Chia-Yi Chen, Shi-Yong Su, Chien-Hsing Chow, Hsiang-Feng Hsiu and Yu-Chine Wang
- [11] SZETO, A. Y. J., and CHRISTIANSEN, K. M.: 'Technological device for deaf-blinded children: need and potential impact', *IEEE Engineering in Medicine and Biology Magazine*, September 1988, pp.25-29
- [12] Telecommunications Standardization Sector of ITLL *Dolo Communication Over the Telephone Network - 600N200-Boud Modem Srondmiked for the General Switched Telephone Nefwork*, ITL Recommendation V.23.
- [13] Independent Communications Authority of South Africa, *Functional requirements for ortologue colling line identifcorion (CLI) customer pmmisses equipment - ICASA TE-010*, May 2002.
- [14] K. R. Ingham, "Braille, the Language, Its Machine Translation and Display", *IEEE Transactions on Publication*, Vol. 10, no. 4, Dec. 1969, pp. 96-100.
- [15] Ying-Wen Bai ,," An adjustable design for the Real Time Clock of high-end server systems" *Electrical and Computer Engineering*, 2008. CCECE 2008. Canadian Conference on,IEEE.
- [16] Sadeque Reza khan, Alvir Kabir, Dilshad Ara Hossain ,,"Designing Smart Multipurpose Digital Clock Using Real Time Clock(RTC) and PIC Microcontoller , *IICA journal* ,volumew41-Number 9,2012.

Stability and Computational Flow Analysis on Boat Hull

A. Srinivas¹, V. Chandra sekhar², Syed Altaf Hussain³

*(PG student, School of Mechanical Engineering RGM College of Engg. & Technology, Nandyal-518501, India)

** (School of Mechanical Engineering, RGM College of Engineering & Technology, Nandyal-518501, India)

*** (School of Mechanical Engineering, RGM College of Engineering & Technology, Nandyal-518501, India)

Abstract : A boat is a water craft of any size designed to float on water to provide passage across water. Boats are extensively used for short distance transportation, trading and fishing etc. This paper describes modeling and CFD (computational fluid dynamics) analysis of boat hull. In the present work, Metacenter height and stability was calculated for a boat hull made of three different materials like wood, aluminum and S-glass/Epoxy composite. The composite materials are being extensively used in varieties of engineering applications because of their superior properties over other conventional materials, the advantages include high strength to weight ratio, high specific strength, high corrosion resistance and high fracture toughness etc. In order to reduce the weight and to increase the performance S-glass/Epoxy composite are used for the boat hull. Theoretical calculation of metacenter height and stability for boat hull is calculated using Simpson's $\frac{1}{3}$ rd rule for the three different boat materials. Modeling and CFD analysis was done for the boat hull made of S-glass/Epoxy composite. To simplify the analysis wind force opposite to the boat hull was taken in to consideration. Flow characteristics like drag and lift forces for different wind and boat speeds were obtained by using Computational Fluid Dynamics (CFD) software by keeping the water speed constant at 5 Km/h. The boat hull geometry is modeled using PRO – E 4.0 and meshing was done using GAMBIT software. Analysis is carried out by using a commercial CFD package of FLUENT.

Keywords – Boat hull, Metacenter, CFD, Aluminum, S-glass/Epoxy composite, Wood.

I. INTRODUCTION

Boat is generally used in sea and ponds. The boats are used for short-distance transportation of people and goods and also for fishing. However many boats are used exclusively for sports. Boats can be made in different shapes and sizes.

They are able to float on water and move due to various propulsion systems (engines, oars, paddles and sails). Boats has been a long fascination, that how to move a boat faster and more safe on the water. Boats have developed from simple rafts to enormous ships that can carry thousands of people.

Boats are made of materials like wood, steel and metal. There are five contributing factors that help to keep a boat afloat and prevent it from sinking. They are buoyancy, stability, waterproofing, air capacity and shape. Buoyancy is an object's ability to float in water. When an object floats in a liquid, the weight of the volume of liquid displaced by the object is equal to the weight of the object.

The stability of a boat ensures that it does not tip over or become unstable. This means all passengers are able to stay in the boat and not get wet.

The boat shape, pointed at one end and wider at the other, is also a design element that helps to move the boat through the water at a faster rate. At the pointed end of boat water and air resistance is minimum. The schematic diagram of boat hull shown in figure: 3.

II. SPECIFICATION OF THE PROBLEM

The objective of the present work is to determine metacenter height and stability of a boat hull made of three different materials viz., Wood, Aluminum and S-glass/Epoxy composite. The structure of the hull varies depending on the vessel type. The material commonly used for the manufacturing boat is steel or wood. If steel is used as the material then weight of the boat increases and wood was less weight and effective corrosion resistance. Modeling and CFD analysis was done for the boat hull made of S-glass/Epoxy composite. 3D model of boat hull geometry is modeled using PRO – E 4.0 and mesh is done using GAMBIT software. Analysis is carried out by using a commercial CFD package of FLUENT. Stability, metacenter height calculations have done by using Simpsons $\frac{1}{3}$ rd rule. The 3D view of a boat hull is show in figure:1 and its specifications are listed in table:1.

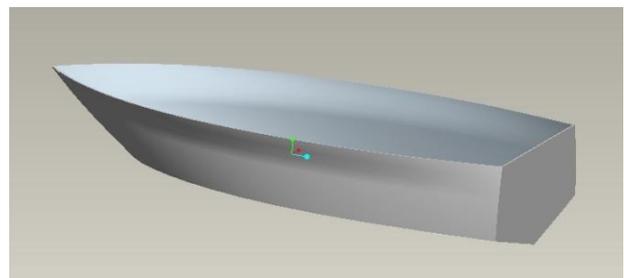


Figure1: 3D view of Boat hull.

III. SPECIFICATION OF BOAT HULL

Table1: Specification of boat hull.

S.NO	PERAMETERS	VALUE (m)
1	Boat hull length	6
2	Boat hull height	2.5
3	Beam	1.2
4	Draft	0.30 (Assumed)
5	Material	Aluminum, Wood and S-Glass/Epoxy composite.

IV. COMPOSITE MATERIAL

Composite materials are formed by combining two or more materials that have quite different properties. The different materials work together to give the composite unique properties. Most composites are made up of just two materials. One material, which is called matrix or binder, surrounds and binds together a cluster of fibers or fragments of a much stronger material, which is the reinforcement. Humans have been using composite materials for thousands of years. Composite materials have extraordinary strength, stiffness, chemical and temperature resistance.

1. E-glass (electrical) - lower alkali content and stronger than A glass (alkali). Good tensile and compressive strength and stiffness, good electrical properties and relatively low cost, but impact resistance relatively poor.
2. C-glass (chemical) - best resistance to chemical attack. Mainly used in the form of surface tissue in the outer layer of laminates used in chemical and water pipes and tanks.
3. R, S or T-glass – manufacturers trade names for equivalent fibers having higher tensile strength and modulus than E glass, with better wet strength retention. Developed for aerospace and defense industries.

V. THEORITICAL CALCULATIONS FOR METACENTER HEIGHT

Simpson’s $\frac{1}{3}$ rd rule may be used to find the areas and volumes of irregular objects. When applied to ships they give a good approximation of areas and volumes. The accuracy of the answers obtained will depend upon the spacing of the ordinates.

This rule assumes that the curve is a parabola of the second order. A parabola of the second order is one whose equation, referred to co-ordinate axis, is of the form $y = a_0 + a_1y + a_2y^2$, where a_0, a_1 and a_2 are constants.

The coefficients of the ordinates are referred to as Simpson’s Multipliers (SM) and they are in the form: 1424241. There had been nine ordinates, the multipliers would have been: 142424241. Simpson’s multiplier begin and end with 1.

Let the curve in Figure:2 be a parabola of the second order. Let y_1, y_2, y_3 and y_4 , represents four ordinates equally spaced at ‘h’ units apart.

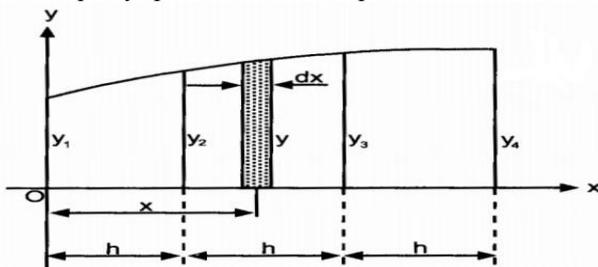


Figure:2 Parabola of the fourth order.

Area of the figure = $\frac{h}{3} (y_1 + y_2 + y_3 + y_4)$

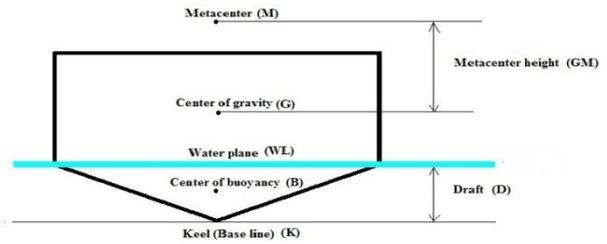


Figure3: Skematic diagram of boat hull.

V.I. AREAS BY USING SIMPSONS RULE

V.1.1 Area of the water plane at 0.30 m draft.

For the purpose of calculation of area of water plane half of the water plane (wp) is considered at the draft of 0.30m. There are six ordinates($y_1, y_2, y_3, y_4, y_5, y_6$) were taken and multiplied with simpson multipliers (SM). Table 2 shows the area of the water plane at various ordinates at 0.30m draft.

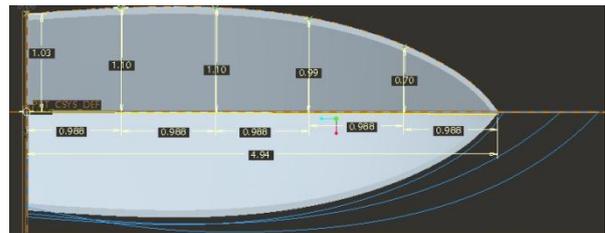


Fig 4: 1/2Area of the water plane at 0.30 m draft

Table 2: 1/2Area of the water plane at 0.30 m draft.

1/2 Ordinate (y)	SM	1/2 Area Of Water plane
1.03	1	1.03
1.10	4	4.4
1.10	2	2.2
0.99	4	3.96
0.70	2	1.4
0	1	0
		$\Sigma = 12.99 \text{ m}$

Common interval (h) = 0.988

1/2 area of water plane = $\frac{1}{3} * h * \Sigma$
 = $\frac{1}{3} * 0.988 * 12.99$
 = 4.278 m^2

Area of the water plane (ΣA) = $\frac{1}{3} * h * \Sigma * 2$
 = 8.566 m^2

V.1.2 Area of the water plane at 0.20 m draft.

For the purpose of calculation of area of water plane half of the water plane (wp) is considered at the draft of 0.20m. There are six ordinates($y_1, y_2, y_3, y_4, y_5, y_6$) were taken and multiplied with simpson multipliers (SM). Table 3 shows the area of the water plane at various ordinate at 0.20m draft.

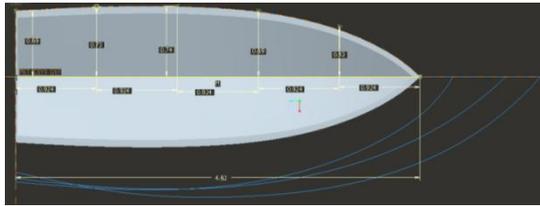


Fig 5: 1/2Area of the water plane at 0.20 m draft

Table 3: 1/2Area of the water plan at 0.20 m draft.

1/2 Ordinate	SM	1/2 Area Of Water plane
0.69	1	0.69
0.73	4	2.92
0.74	2	1.48
0.60	4	2.4
0.53	2	1.06
0	1	0
		$\Sigma = 8.55 \text{ m}$

Common interval (h) = 0.924

$$\begin{aligned} \frac{1}{2} \text{ area of water plane} &= \frac{1}{3} * h * \Sigma \\ &= \frac{1}{3} * 0.924 * 8.55 \\ &= 2.633 \text{ m}^2 \end{aligned}$$

$$\begin{aligned} \text{Area of the water plane } (\Sigma A) &= \frac{1}{3} * h * \Sigma * 2 \\ &= 5.26 \text{ m}^2 \end{aligned}$$

V.1.3 Area of the water plane at 0.10 m draft.

For the purpose of calculation of area of water plane half of the water plane (wp) is considered at the draft of 0.10m. There are six ordinates ($y_1, y_2, y_3, y_4, y_5, y_6$) were taken and multiplied with simpson multipliers(SM). Table 4 shows the area of the water plane at various ordinate at 0.10m draft.

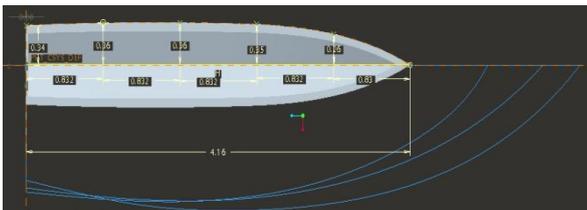


Fig 6: 1/2Area of the water plane at 0.10 m draft.

Table 4: 1/2Area of the water plane at 0.10 m draft

1/2 Ordinate	SM	1/2 Area Of Water plane
0.34	1	0.34
0.36	4	1.44
0.36	2	0.72
0.35	4	1.4
0.26	2	0.52
0	1	0
		$\Sigma = 4.42 \text{ m}$

Common interval (h) = 0.832

$$\begin{aligned} \frac{1}{2} \text{ area of water plane} &= \frac{1}{3} * h * \Sigma (\frac{1}{2} \text{ area}) \\ &= \frac{1}{3} * 0.832 * 4.42 \\ &= 1.22 \text{ m}^2 \end{aligned}$$

$$\begin{aligned} \text{Area of the water plane } (\Sigma A) &= \frac{1}{3} * h * \Sigma * 2 \\ &= 2.44 \text{ m}^2 \end{aligned}$$

V.2. VOLUME OF WATER DISPLACED:

For the purpose of calculation of volume of water displaced of boat hull at the draft of 0.30m for different materials.

V.2.1 For S-glass/Epoxy composite:- (density = 2000 Kg/m³):

$$\begin{aligned} \text{Weight of boat hull} &= \text{volume of boat hull} * \text{density of boat hull material} * g \\ &= 1.077 * 2000 * 9.81 \\ &= 21130.74 \text{ N} \end{aligned}$$

For equilibrium weight of water displaced = weight of boat hull

$$\begin{aligned} \text{Volume of water displaced} &= \frac{\text{weight of water displaced}}{\text{water density}} \\ &= \frac{21130.74}{(1025 * 9.81)} \\ &= 2.10 \text{ m}^3 \end{aligned}$$

V.2.2 Wood :- (density = 700 Kg/m³)

$$\begin{aligned} \text{Weight of boat hull} &= \text{volume of boat hull} * \text{density of boat hull material} * g \\ &= 1.077 * 700 * 9.81 \\ &= 7395.75 \text{ N} \end{aligned}$$

For equilibrium weight of water displaced = weight of boat hull

$$\begin{aligned} \text{Volume of water displaced} &= \frac{\text{weight of water displaced}}{\text{water density}} \\ &= \frac{7395.75}{(1025 * 9.81)} \\ &= 0.735 \text{ m}^3 \end{aligned}$$

V.2.3 Aluminum :- (density = 2730 Kg/m³)

$$\begin{aligned} \text{Weight of boat hull} &= \text{volume of boat hull} * \text{density of boat hull material} * g \\ &= 1.077 * 2730 * 9.81 \\ &= 28843.46 \text{ N} \end{aligned}$$

For equilibrium weight of water displaced = weight of boat hull

$$\begin{aligned} \text{Volume of water displaced} &= \frac{\text{weight of water displaced}}{\text{water density}} \\ &= \frac{28843.46}{(1025 * 9.81)} \\ &= 2.86 \text{ m}^3 \end{aligned}$$

V.3 CENTER OF BUOYANCY:

For the purpose of calculation of center of buoyancy at the draft of 0.30m the area of the half of the water plane multiplied with simpson multipliers (SM). Table 5 shows the center of buoyancy at the draft of 0.30m.

Table 5: Center of buoyancy.

Area of wp	SM	Product
1.03	1	1.03
4.4	4	17.6
2.2	2	4.4
3.96	4	15.84
1.4	2	2.8
0	1	0
$\Sigma = 12.99 \text{ m}^2$		$\Sigma = 41.67 \text{ m}^2$

$$CB = \frac{A_{WP} (SM) Z}{\text{product}}$$

Where,

A_{WP} = water plane area Z = vertical distance b/w the base line to water plane

$$CB = \frac{(12.99 * 0.30)}{41.67} = 0.09$$

V.4 SECOND MOMENT OF THE WATER LINE AREA ABOUT THE CENTERLINE

For the purpose of calculation of second moment of water line area about the centerline six ordinates ($y_1, y_2, y_3, y_4, y_5, y_6$) were taken and their cubic value was multiplied with Simpson multipliers (SM). Table 6 shows the second moment of the water line area about the centerline.

Table 6: Second moment of the water line area.

1/2 ordinate	(1/2 ordinate) ³	SM	Product
1.03	1.092	1	1.092
1.10	1.33	4	5.32
1.10	1.33	2	2.66
0.99	0.97	4	3.88
0.70	0.34	2	0.68
0	0	1	0
			$\sum = 13.632$ m^3

$$I_{cl} = \frac{2}{9} * h * \text{product} = \frac{2}{9} * 0.988 * 13.632 = 2.99 m^4$$

V.5 TO FIND THE DISTANCE OF “ BM “ :

$$BM = \frac{I}{V}$$

I = Second moment of the water plane area about the centerline

V = The volume of water displacement

V.5.1 For S-glass/Epoxy composite:

$$BM = \frac{2.99}{2.10} = 1.42 m$$

Distance of “ KM “ :

$$KM = KB + BM = 0.09 + 1.42 = 1.51 m$$

V.5.2 For Aluminium:

$$BM = \frac{2.99}{2.86} = 1.045 m$$

Distance of “ KM “ :

$$KM = KB + BM = 0.09 + 1.045 = 1.135 m$$

V.5.3 For Wood:

$$BM = \frac{2.99}{0.735} = 4.06 m$$

Distance of “ KM “ :

$$KM = KB + BM = 0.09 + 4.06 = 4.15 m$$

V.6 STABILITY IN THE UPRIGHT CONDITIONS

V.6.1 For S-glass/Epoxy composite:

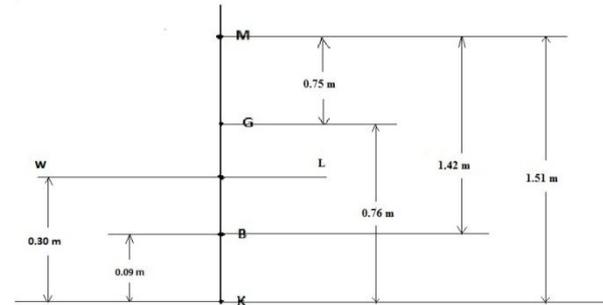


Fig 7: Stability in the upright conditions for S-glass/epoxy.

V.6.2 For Aluminium:

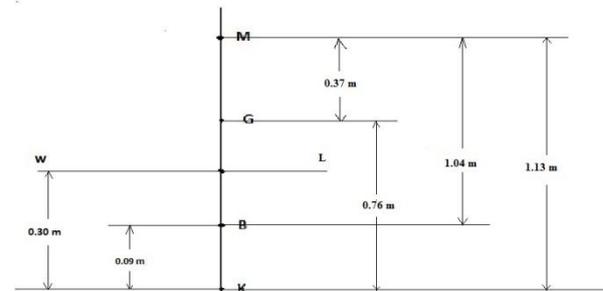


Fig 8: Stability in the upright conditions aluminium.

V.6.3 For wood:

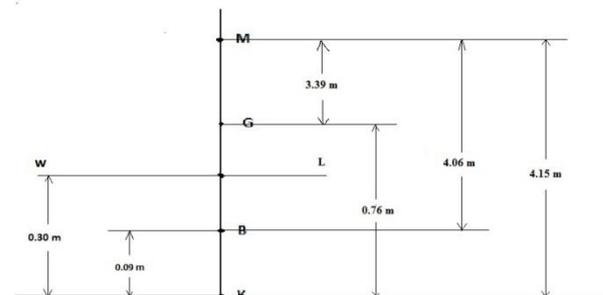


Fig 9: Stability in the upright conditions wood.

1. If G is below M, the boat in stable equilibrium.
2. If G is above M, the boat in unstable equilibrium.
3. If G is coincide with M, the boat is in neutral.

VI. DRAG AND LIFT FORCES ACTING ON BOAT

Drag is the force parallel to the relative wind.

Lift is the force perpendicular to the relative wind.

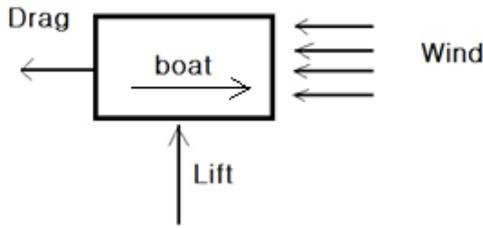


Fig10: Drag and Lift forces acting on boat.

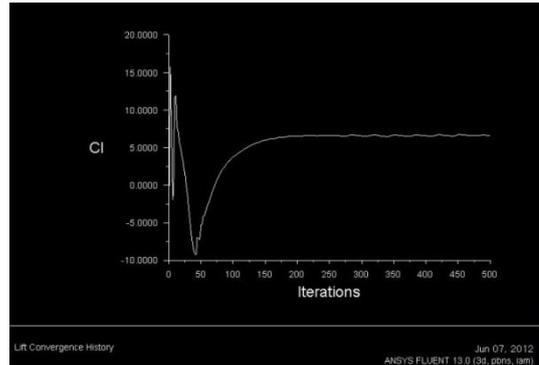


Fig14: Lift History.

VII. COMPUTATIONAL FLUID DYNAMICS (CFD) RESULTS

VII.1 Analysis of Boat with speed of 35 Km/h, Wind speed as 20 Km/h, by keeping Water speed as 5 Km/h. Static pressure, velocity, drag, lift history as shown below

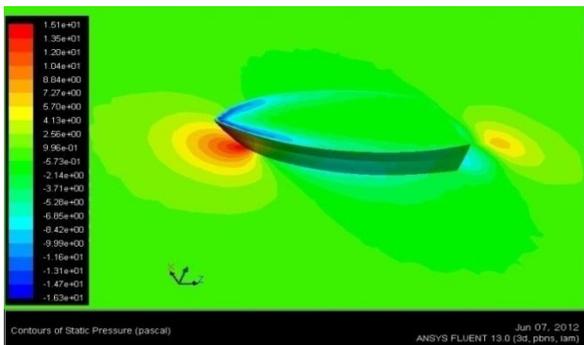


Fig11: Static Pressure on boat hull.

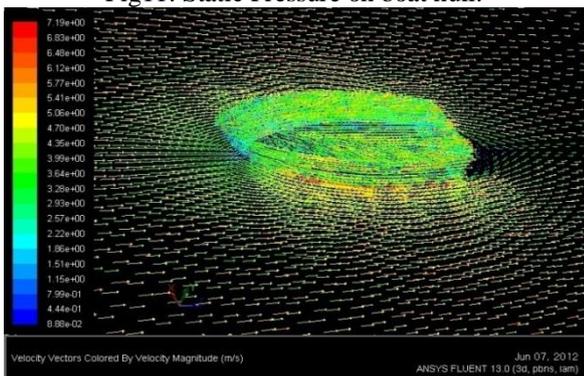


Fig12: Velocity on boat hull.

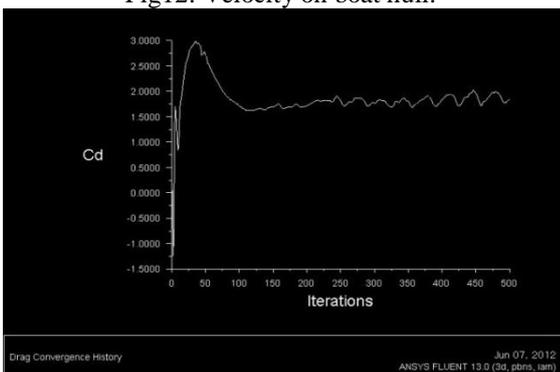


Fig13: Drag History.

VIII. THEORITICAL FORMULAS FOR DRAG AND LIFT FORCES

$$\text{Drag force} = \frac{1}{2} * \rho * V^2 * C_D * A$$

$$\text{Lift force} = \frac{1}{2} * \rho * V^2 * C_L * A$$

Where ρ = Density of water
 V = Water velocity in m/s
 C_D = Drag coefficient
 C_L = Lift coefficient
 A = Area of the boat (28.633 m²)

Similar analysis was carried out for different wind speeds (30, 40 and 50 Km/h) and boat speeds (45, 55 and 65 Km/h), by keeping water speed as 5 Km/h. Results for Drag, Lift coefficients are tabulated in tables from 7 to 10 and the respective drag and lift forces are tabulated in tables 11 to 14

IX. CFD RESULTS FOR DRAG AND LIFT COEFFICIENTS

Drag and lift coefficients are vary depend upon boat speed.

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Coefficient	Lift Coefficient
1	35	5	20	1.84	6.58
2	45	5	20	1.91	6.57
3	55	5	20	1.92	6.62
4	65	5	20	2.01	6.56

Table:7

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Coefficient	Lift Coefficient
1	35	5	30	4.03	14.99
2	45	5	30	3.99	14.94
3	55	5	30	4.19	14.82
4	65	5	30	4.22	14.79

Table:8

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Coefficient	Lift Coefficient
1	35	5	40	6.79	26.58
2	45	5	40	7.61	27.19
3	55	5	40	7.88	26.94
4	65	5	40	6.87	26.25

Table:9

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Coefficient	Lift Coefficient
1	35	5	50	11.29	42.58
2	45	5	50	11.28	41.79
3	55	5	50	11.37	42.23
4	65	5	50	12.05	41.24

Table:10

X. CFD RESULTS FOR DRAG AND LIFT FORCES

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Force (N)	Lift Force (N)
1	35	5	20	5.66	19.46
2	45	5	20	5.45	19.41
3	55	5	20	5.69	19.58
4	65	5	20	5.94	19.41

Table:11

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Force (N)	Lift Force (N)
1	35	5	30	12.43	44.31
2	45	5	30	12.88	44.16
3	55	5	30	12.39	43.82
4	65	5	30	12.21	43.71

Table:12

S.No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Force (N)	Lift Force (N)
1	35	5	40	20.32	78.56
2	45	5	40	23.59	80.37
3	55	5	40	20.12	79.62
4	65	5	40	20.03	77.60

Table:13

S. No	Boat speed (Km/h)	Water speed (Km/h)	Wind speed (Km/h)	Drag Force (N)	Lift Force (N)
1	35	5	50	33.37	125.87
2	45	5	50	33.35	123.52
3	55	5	50	33.63	124.82
4	65	5	50	35.63	121.89

Table:14

XI. CONCLUSIONS

The following conclusion were drawn from the CFD analysis performed on a boat hull of S-glass/Epoxy composite.

For stability, the center of gravity (G) is below the metacenter (M). This condition was validated for the three different materials considered for the analysis.

Minimum drag force of 5.45 N and minimum lift force of 19.41 N was obtained for the boat speed of 45 Km/h, when the water and wind speeds are 5 Km/h, 20 Km/h respectively.

Minimum drag force of 12.21 N and minimum lift force of 43.71 N was obtained for the boat speed of 65 Km/h, when the water and wind speeds are 5 Km/h, 30 Km/h respectively.

Minimum drag force of 20.03 N and minimum lift force of 77.60 N was obtained for the boat speed of 65 Km/h, when the water and wind speeds are 5 Km/h, 40 Km/h respectively.

Minimum drag force of 33.35 N and minimum lift force of 121.89 N was obtained for the boat speed of 45 Km/h and 65 Km/h, when the water and wind speeds are 5 Km/h, 50 Km/h respectively.

Drag and lift forces increases with increase in wind speed as well as with boat speed.

For various wind speeds (20,30,40,50 Km/h). Minimum drag force and Minmum lift force was obtained at 45 Km/h, 65 km/h of boat speeds, when water speed is 5 Km/h.

REFERENCES

- [1] Ship Stability for Masters and Mates, Sixth Edition.
- [2] A Matter of Stability & Trim.
- [3] A TextBook of Fluid Mechanics and Hydraulic Machines - Dr. R. K. Bansal.
- [4] Basic ship theory, vol2.
- [5] Principles Of Naval Architecture Vol I - Stability And Strength.

Power quality improvement of Distribution lines using DSTATCOM under various loading conditions

Veeraiah Kumbha

*M-Tech Scholar, Power systems,
Department Of Electrical And Electronics Engineering,
JNTU KAKINADA (A.P), India.*

N. Sumathi

*Assistant Professor, M. Tech, (Ph. D) High Voltage
Engineering, Department of Electrical and Electronics
Engineering, JNTU KAKINADA (A.P), India.*

Abstract: A Power quality problem is an occurrence manifested as a nonstandard voltage, current or frequency that results in a failure or a mis-operation of end user equipments. Utility distribution networks, sensitive industrial loads and critical commercial operations suffer from various types of outages and service interruptions which can cost significant financial losses. With the restructuring of power systems and with shifting trend towards distributed and dispersed generation, the issue of power quality is going to take newer dimensions. Injection of the wind power into an electric grid affects the power quality. The performance of the wind turbine and thereby power quality are determined on the basis of measurements and the norms followed according to the guideline specified in International Electro-technical Commission standard, IEC-61400. The influence of the wind turbine in the grid system concerning the power quality measurements are-the active power, reactive power, variation of voltage, flicker, harmonics, and electrical behavior of switching operation and these are measured according to national/international guidelines. The paper study demonstrates the power quality problem due to installation of wind turbine with the grid. In this proposed scheme STATIC COMPENSATOR (STATCOM) is connected at a point of common coupling with a battery energy storage system (BESS) to mitigate the power quality issues. The battery energy storage is integrated to sustain the real power source under fluctuating wind power. The STATCOM control scheme for the grid connected wind energy generation system for power quality improvement is simulated using MATLAB/SIMULINK in power system block set. Finally the proposed scheme is applied for both balanced and unbalanced linear non linear loads.

Index Terms: International electro-technical commission (IEC), power quality, wind generating system (WGS).

I. Introduction

Electric Power quality is a term which has captured increasing attention in power engineering in the recent years. Eventhough this subject has always been of interest to power engineers, it has assumed considerable interest in the 1990's. Electric power quality means different things for different people. To most electric power engineers, the term refers to a certain sufficiently high grade of electric service but beyond that there is no universal agreement. The measure of power quality depends upon the needs of the equipment that is being supplied. What is good power quality for an electric motor may not be good enough for a personal computer. Usually the term power quality refers to maintaining a sinusoidal waveform of bus voltages at rated voltage and frequency.

The waveform of electric power at generation stage is purely sinusoidal and free from any distortion. Many of the Power conversion and consumption equipment are also designed to function under pure sinusoidal voltage waveforms. However, there are many devices that distort the waveform. These distortions may propagate all over the electrical network. In recent years, there has been an increased use of non-linear loads which has resulted in an increased fraction of non-sinusoidal currents and voltages in Electric Network. Classification of power quality areas may be made according to the source of the problem such as converters, magnetic circuit non linearity, arc furnace or by the wave shape of the signal such as harmonics, flicker or by the frequency spectrum (radio frequency interference). The wave shape phenomena associated with power quality may be characterized into synchronous and nonsynchronous Phenomena. Synchronous phenomena refer to those in synchronism with A.C waveform at power frequency.

The main aspects of electric power quality

May be categorized as:-

- a) Fundamental concepts
- b) Sources
- c) Instrumentation
- d) Modeling
- e) Analysis
- f) Effects

Figure 1 shows some of the typical voltage disturbances.

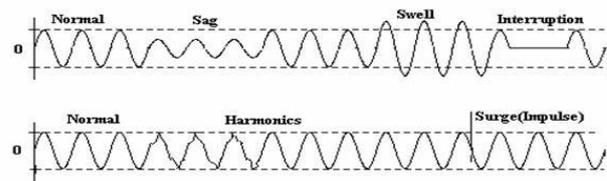


Figure.1 Typical Voltage Disturbances Voltage disturbance as shown in above figure

Table-I

Power Freq Disturbance	Electro Magnetic Interferences	Power System Transient	Power System Harmonics	Electrostatic Discharge	Power Factor
<ul style="list-style-type: none"> • Low Freq phenomena • Produce Voltage sag / swell 	<ul style="list-style-type: none"> • High freq phenomena • interaction between electric and magnetic field 	<ul style="list-style-type: none"> • Fast, short-duration event • Produce distortion like notch, impulse 	<ul style="list-style-type: none"> • Low frequency phenomena • Produce waveform distortion 	<ul style="list-style-type: none"> • Current flow with different potentials • Caused by direct current or induced electrostatic field 	<ul style="list-style-type: none"> • Low power factor causes equipment damage

Table1 shows the different power quality problems

Custom power Devices like STATCOM (shunt active power filter), DVR and UPQC (combination of series and shunt active power filter) are the latest development of interfacing devices between distribution supply and

consumer appliances to overcome voltage/current disturbances and improve the power quality by compensating the reactive and harmonic power generated or absorbed by the load. Wind is the most promising DG sources and their penetration level to the grid is also on the rise. Although the benefits of DG includes voltage support, diversification of power sources, reduction in transmission and distribution losses and improved reliability.

The solutions of STATCOM is often used in transmission system. When it is used in distribution system, it is called D-STATCOM (STATCOM in Distribution system). D-STATCOM is a key FACTS controller and it utilizes power electronics to solve many power quality problems commonly faced by distribution systems. Potential applications of D-STATCOM include power factor correction, voltage regulation, load balancing and harmonic reduction. Comparing with the SVC, the D-STATCOM has quicker response time and compact structure. It is expected that the D-STATCOM will replace the roles of SVC in nearly future D-STATCOM and STATCOM are different in both structure and function, while the choice of control strategy is related to the main-circuit structure and main function of compensators [3], so D-STATCOM and STATCOM adopt different control strategy. At present, the use of STATCOM is wide and its strategy is mature, while the introduction of D-STATCOM is seldom reported. Many control techniques are reported such as instantaneous reactive power theory (Akagi et al., 1984), power balance theory, etc. In this paper, an indirect current control technique (Singh et al., 2000a,b) is employed to obtain gating signals for the Insulated Gate Bipolar Transistor (IGBT) devices used in current controlled voltage source inverter (CC-VSI) working as a DSTATCOM. A model of DSTATCOM is developed using MATLAB for investigating the transient analysis of distribution system under balanced/unbalanced linear and non-linear three-phase and single-phase loads (diode rectifier with R and R-C load). Simulation results during steady-state and transient operating conditions of the DSTATCOM are presented and discussed to demonstrate power factor correction, harmonic elimination and load balancing capabilities of the DSTATCOM system [5-10].

II. WIND ENERGY SYSTEM

A simplified diagram representing some of the common types of wind energy systems are shown in Fig.2. From the design perspective it is found that some generators are directly connected to the grid through a dedicated transformer while others incorporate power electronics. Many designs, however, include some level of power Electronics to improve controllability and operating range. Whatever connection configuration is used, each turbine itself has an effect on the power quality of the transmission system. Recent analysis and study shows that the impact of the yaw error and horizontal wind shear on the power (torque) and voltage oscillations is more severe than the effects due to the tower shadow and vertical wind shear.

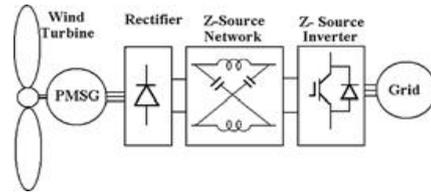


Fig.2 Different types of wind energy system
The above figure shows different types of wind energy system

The new grid comes adopted for wind power integration has identified the problems of integrating large amounts of wind energy to the electric grid. It suggests that new wind farms must be able to provide voltage and reactive power control, frequency control and fault ride-through capability in order to maintain the electric system stability. For the existing wind farms with variable speed, double-fed induction generators (DFIG) and synchronous generators (SG), a frequency response in the turbine control system can be frequency response in the turbine control system can be incorporated by a software upgrade. Wind farms with fixed speed induction generators (FSIG) have to be phased out because they cannot offer the required voltage or frequency control. An overview of the developed controllers for the converter of grid connected system and showed that the DFIG has now the most efficient design for the regulation of reactive power and the adjustment of angular velocity to maximize the output power efficiency. These generators can also support the system during voltage sags. However, the drawbacks of converter-base systems are harmonic distortions injected into the system. Being a single-stage buck-boost inverter, with Z-source inverter can be a good candidate to mitigate the PQ problems for future DG systems connected to the grid Fig(3).

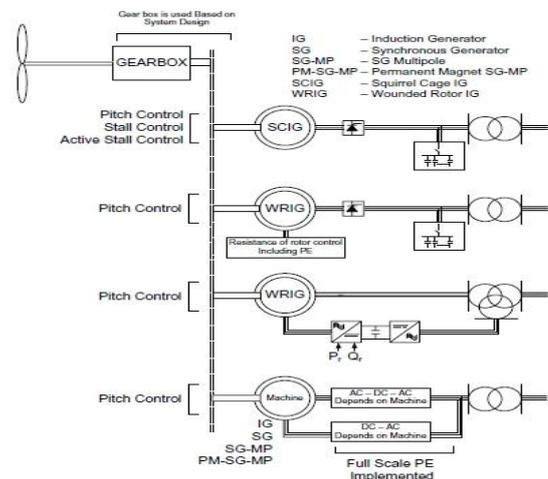


Fig.3 PMSG-base WECS with dc boost chopper and ZSI

Anti-islanding is one of the important issues for grid connected DG system. A major challenge for the islanding operation and control schemes is the protection coordination of distribution systems with bidirectional flows of fault current. This is unlike the conventional over-current protection for radial systems with unidirectional flow of fault current. Therefore extensive research is being carried out and an overview of the existing protection techniques with islanding operation and control, for preventing disconnection of DGs during loss of grid, has been discussed.

III. MITIGATION OF PQ PROBLEMS

There are two ways to mitigate the power quality problems-either from the customer side or from the utility side. The first approach is called load conditioning, which ensures that the equipment is less sensitive to power disturbances, allowing the operation even under significant voltage distortion. The other solution is to install line conditioning systems that suppress or counteracts the power system disturbances. Several devices including flywheels, super-capacitors, other energy storage systems, constant voltage transformers, noise filters, isolations transformers, transient voltage surge suppressors, harmonic filters are used for the mitigation of specific PQ problems. Custom power devices (CPD) like DSTATCOM, DVR and UPQC are capable of mitigating multiple PQ problems associated with utility distribution and the end used appliance. The following section looks at the role of CPDs in mitigating PQ problems in relation to grid integrated with wind energy systems.

IV. Distribution Static Compensator (DSTATCOM)

5.1 System configuration DSTATCOM

DSTATCOM is a shunt-connected custom power device specially designed for power factor correction, current harmonics filtering and load balancing. It can also be used for voltage regulation at a distribution bus. It is often referred to as a shunt or parallel active power filter. It consists of a voltage or a current source PWM converter Fig.4. It operates as a current controlled voltage source and compensates current harmonics by injecting the harmonic components generated by the load but phase shifted by 180 degrees. With an appropriate control scheme, the DSTATCOM can also compensate for poor load power factor.

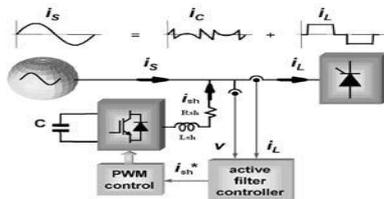


Fig.4 system configuration of DSTATCOM

The system configuration of DSTATCOM above shows figure

When the STATCOM is applied in distribution system is called DSTATCOM (Distribution-STATCOM) and its configuration is the same, or with small modifications, oriented to a possible future amplification of its possibilities in the distribution network at low and medium voltage, implementing the function so that we can describe as flicker damping, harmonic filtering and hole and short interruption compensation.

Distribution STATCOM (DSTATCOM) exhibits high speed control of reactive power to provide voltage stabilization, flicker suppression, and other types of system control. The DSTATCOM utilizes a design consisting of a GTO- or IGBT-based voltage sourced converter connected to the power system via a multi-stage converter transformer.

The DSTATCOM protects the utility transmission or distribution system from voltage sags and/or flicker caused by rapidly varying reactive current demand. In utility applications, a DSTATCOM provides leading or lagging reactive power to achieve system stability during transient conditions.

The DSTATCOM can also be applied to industrial facilities to compensate for voltage sag and flicker caused by non-linear dynamic loads, enabling such problem loads to co-exist on the same feeder as more sensitive loads. The DSTATCOM instantaneously exchanges reactive power with the distribution system without the use of bulky capacitors or reactors.

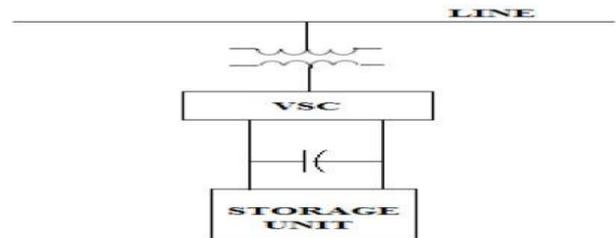


Fig.4 Basic circuit diagram of DSTATCOM

The above figure shows basic circuit diagram of distribution static synchronous compensator.

In most applications, a DSTATCOM can use its significant short-term transient overload capabilities to reduce the size of the compensation system needed to handle transient events. The short-term overload capability is up to 325% for periods of 1 to 3 seconds, which allows applications such as wind farms and utility voltage stabilization to optimize the system's cost and performance. The DSTATCOM controls traditional mechanically switched capacitors to provide optimal compensation on a both a transient and steady state basis. To prevent the unbalanced and distorted currents from being drawn from the distribution bus, a shunt compensator, DSTATCOM, can be used to ensure that the current drawn from the distribution bus is balanced and sinusoidal. A Voltage Source Converter (VSC) is used to realize a DSTATCOM. The structure of the VSC decides the extent of compensation it can provide.

5.2 Voltage Source Converter (VSC)

A voltage-source converter is a power electronic device that connected in shunt or parallel to the system. It can generate a sinusoidal voltage with any required magnitude, frequency and phase angle. The VSC used to either completely replace the voltage or to inject the 'missing voltage'. The 'missing voltage' is the difference between the nominal voltage and the actual. It also converts the DC voltage across storage devices into a set of three phase AC output voltages [8, 9]. In addition, D-STATCOM is also capable to generate or absorbs reactive power. If the output voltage of the VSC is greater than AC bus terminal voltages, D-STATCOM is said to be in capacitive mode. So, it will compensate the reactive power through AC system and regulates missing voltages. These voltages are in phase and coupled with the AC system through the reactance of coupling transformers. Suitable adjustment of the phase and magnitude of the DSTATCOM output voltages allows effective control of active and reactive

power exchanges between D-STATCOM and AC system. In addition, the converter is normally based on some kind of energy storage, which will supply the converter with a DC voltage [10].

5.3 Controller for DSTATCOM

The three-phase reference source currents are computed using three-phase AC voltages (v_{ta}, v_{tb} and v_{tc}) and DC bus voltage (V_{dc}) of DSTATCOM. These reference supply currents consist of two components, one in-phase (I_{spdr}) and another in quadrature (I_{spqr}) with the supply voltages. The control scheme is represented in Fig. 5. The basic equations of control algorithm of DSTATCOM are as follows.

5.3.1 Computation of in-phase components of reference supply current

The instantaneous values of in-phase component of reference supply currents (I_{spdr}) is computed using one PI controller over the average value of DC bus voltage of the DSTATCOM (v_{dc}) and reference DC voltage (v_{dcr}) as

$$I_{spdr} = I_{spdr(n-1)} + K_{pd} \{V_{dc(n)} - V_{dc(n-1)}\} + K_{id} V_{dc(n)}$$

where V_{dc}(n) □ V_{dcc}-V_{dcn} denotes the error in V_{dcc} and average value of V_{dc} K_{pd} and K_{id} are proportional and integral gains of the DC bus voltage PI controller. The output of this PI controller (I_{spdr}) is taken as amplitude of in-phase component of the reference supply currents. Three-phase in-phase components of the reference supply currents (i_{sadr}, i_{sbr} and i_{sdr}) are computed using the in-phase unit current vectors (u_a, u_b and u_c) derived from the AC terminal voltages (v_{tan}, v_{tb} and v_{tc}), respectively.

$$U_a = V_{ta}/V_{tm} \quad U_b = V_{tb}/V_{tm} \quad U_c = V_{tc}/V_{tm}$$

where V_{tm} is amplitude of the supply voltage and it is computed as

$$V_{tm} = \sqrt{[(2/3)(V_{tan}^2 + V_{tb}^2 + V_{tc}^2)]}$$

The instantaneous values of in-phase component of reference supply currents (i_{sadr}, i_{sbr} and i_{sdr}) are computed as

$$I_{sadr} = I_{spds} U_a \quad I_{sbr} = I_{spdr} U_b \quad I_{sdr} = I_{spdr} U_c$$

5.3.2 Computation of quadrature components of reference supply current

The amplitude of quadrature component of reference supply currents is computed using a second PI controller over the amplitude of supply voltage (v_{tm}) and its reference value (v_{tmr})

$$I_{spqr}(n) = I_{spqr}(n-1) + K_{pq} \{V_{ac}(n) - V_{ac}(n-1)\} + K_{iq} V_{ac}(n)$$

Where V_{ac} = V_{tm}-V_{mc}(n) denotes the error in V_{tm} and computed value V_{tm} from Equation (3) and K_{pq} and K_{iq} are the proportional and integral gains of the second PI controller.

$$W_a = \{-U_b + U_c\}/\{\sqrt{3}\}$$

$$W_b = \{U_a\sqrt{3} + U_b - U_c\}/\{2\sqrt{3}\}$$

$$W_c = \{-U_a\sqrt{3} + U_b - U_c\}/\{2\sqrt{3}\}$$

Three-phase quadrature components of the reference supply currents (i_{saqr}, i_{sbqr} and i_{scqr}) are computed using the output of second PI controller (I_{spqr}) and quadrature unit current vectors (w_a, w_b and w_c) as

$$i_{saqr} = I_{spqr} W_a, \quad i_{sbqr} = I_{spqr} W_b, \quad i_{scqr} = I_{spqr} W_c,$$

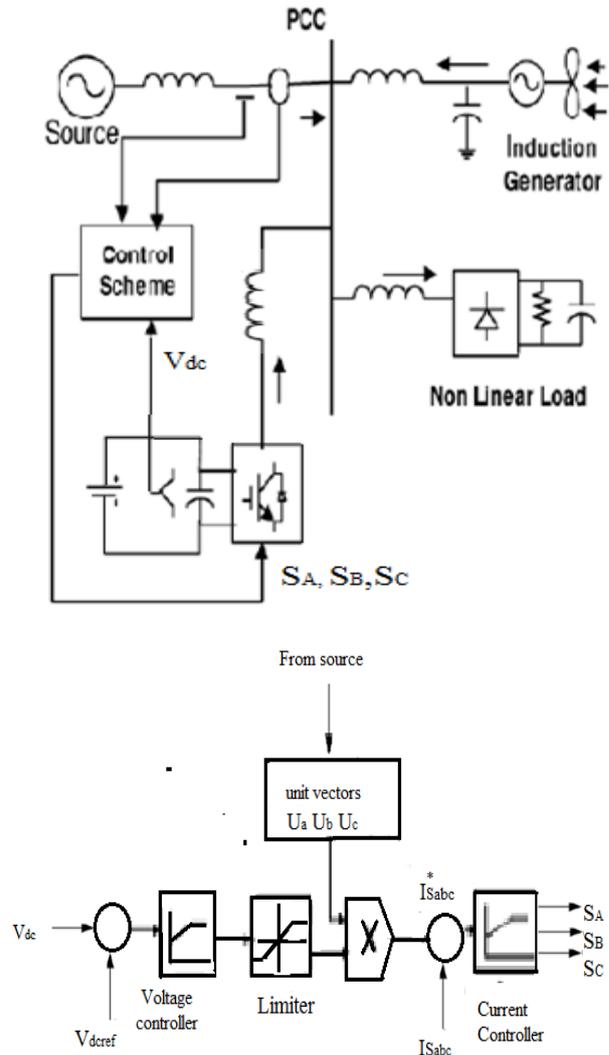


Figure. 5 Control methods for DSTATCOM

5.3.3 Computation of total reference supply currents

Three-phase instantaneous reference supply currents (i_{sar}, i_{sbr} and i_{sdr}) are computed by adding in-phase (i_{sadr}, i_{sbr} and i_{sdr}) and quadrature components of supply currents (i_{saqr}, i_{sbqr} and i_{scqr}) as

$$i_{sar} = i_{sadr} + i_{saqr}, \quad i_{sbr} = i_{sbr} + i_{sbqr},$$

$$i_{sdr} = i_{sdr} + i_{scqr}$$

A hysteresis pulse width modulated (PWM) current controller is employed over the reference (i_{sar}, i_{sbr} and i_{sdr}) and sensed supply currents (i_{sa}, i_{sb} and i_{sc}) to generate gating pulses for IGBTs of DSTATCOM.

V. MATAB/SIMULINK MODELING OF DSTATCOM

6.1 Modeling of Power Circuit

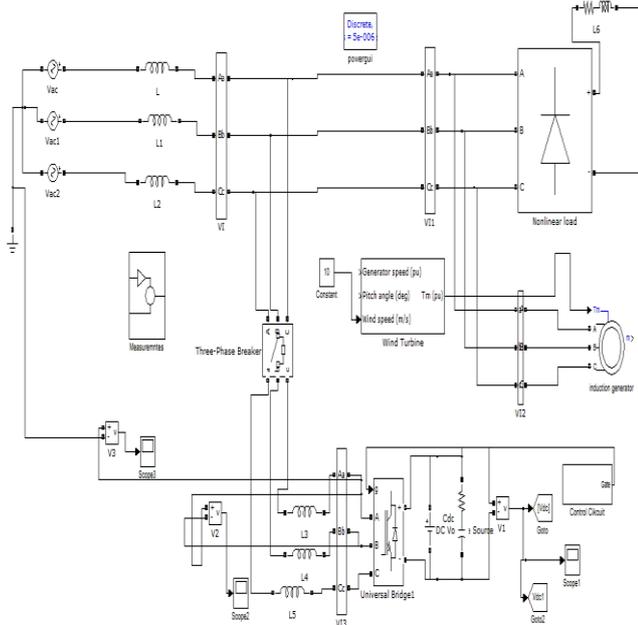


Figure. 6 Matlab/Simulink Model of DSTATCOM

Fig. 6 shows the complete MATLAB model of DSTATCOM along with control circuit. The power circuit as well as control system are modelled using Power System Blockset and Simulink. The grid source is represented by three-phase AC source. Three-phase AC loads are connected at the load end. DSTATCOM is connected in shunt and it consists of PWM voltage source inverter circuit and a DC capacitor connected at its DC bus. An IGBT-based PWM inverter is implemented using Universal bridge block from Power Electronics subset of PSB. Snubber circuits are connected in parallel with each IGBT for protection. Simulation of DSTATCOM system is carried out for linear and non-linear loads. The linear load on the system is modelled using the block three-phase parallel R-L load connected in delta configuration. The non-linear load on the system is modelled using R and R-C circuits connected at output of the diode rectifier. Provision is made to connect loads in parallel so that the effect of sudden load addition and removal is studied. The feeder connected from the three-phase source to load is modelled using appropriate values of resistive and inductive components.

5.1 Modeling of Control Circuit

Fig. 7 shows the control algorithm of DSTATCOM with two PI controllers. One PI controller regulates the DC link voltage while the second PI controller regulates the terminal voltage at PCC. The in-phase components of DSTATCOM reference currents are responsible for power factor correction of load and the quadrature components of supply reference currents are to regulate the AC system voltage at PCC.

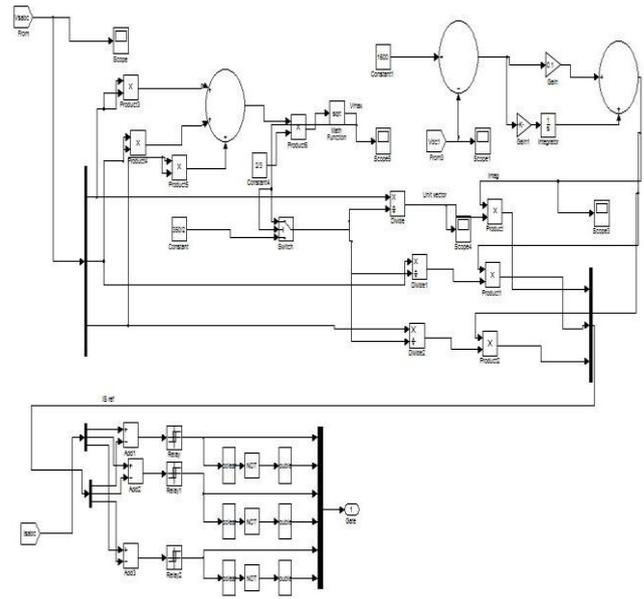


Figure. 7 Control Circuit

The output of PI controller over the DC bus voltage (I_{spdr}) is considered as the amplitude of the in-phase component of supply reference currents and the output of PI controller over AC terminal voltage (I_{spqr}) is considered as the amplitude of the quadrature component of supply reference currents. The instantaneous reference currents (i_{sar} , i_{sbr} and i_{scr}) are obtained by adding the in-phase supply reference currents (i_{sadr} , $i_{s bdr}$ and $i_{s cdr}$) and quadrature supply reference currents (i_{saqr} , i_{sbqr} and i_{scqr}). Once the reference supply currents are generated, a carrierless hysteresis PWM controller is employed over the sensed supply currents (i_{sa} , i_{sb} and i_{sc}) and instantaneous reference currents (i_{sar} , i_{sbr} and i_{scr}) to generate gating pulses to the IGBTs of DSTATCOM. The controller controls the DSTATCOM currents to maintain supply currents in a band around the desired reference current values. The hysteresis controller generates appropriate switching pulses for six IGBTs of the VSI working as DSTATCOM.

VI. SIMULATION RESULTS

Here Simulation results are presented for four cases. In case one load is balanced non linear, case two load is unbalanced non linear, case three load is balanced linear and in case four unbalanced linear load is considered.

6.1 Case one

Performance of DSTATCOM connected to a weak supply system is shown in Fig.5 for power factor correction and load balancing. This figure shows variation of performance variables such as supply voltages (v_{sa} , v_{sb} and v_{sc}), terminal voltages at PCC (v_{ta} , v_{tb} and v_{tc}), supply currents (i_{sa} , i_{sb} and i_{sc}), load currents (i_{la} , i_{lb} and i_{lc}), DSTATCOM currents (i_{ca} , i_{cb} and i_{cc}) and DC link voltage (V_{dc}).

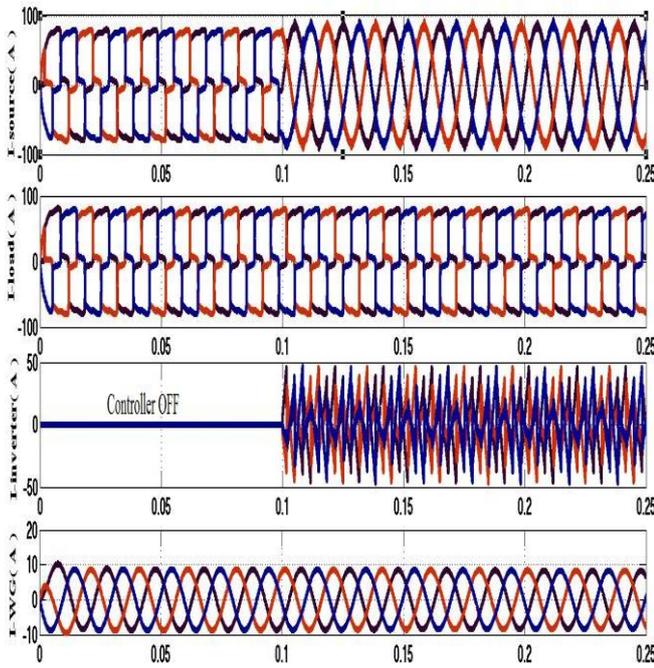


Figure. 8 Simulation results for Balanced Non Linear Load
 (a) Source current. (b) Load current. (c) Inverter injected current.
 (d) wind generator (induction generator) current.

Fig. 8 shows the source current, load current and compensator current and induction generator currents plots respectively. Here compensator is turned on at 0.1 seconds.

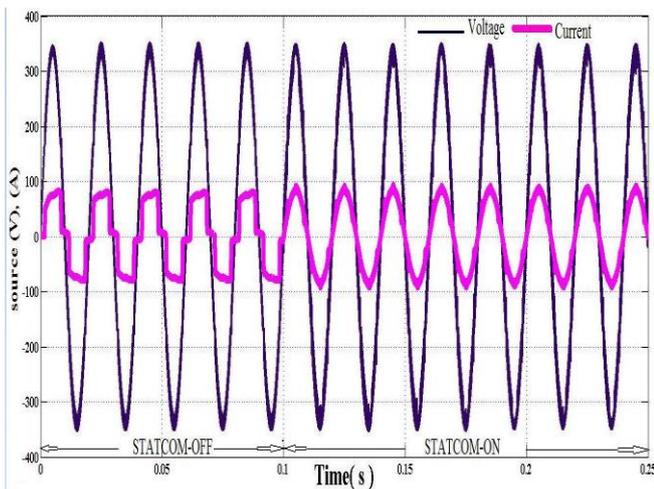


Figure. 9 Simulation results power factor for Non linear Load

Fig. 9 show the power factor it is clear from the figure after compensation power factor is unity.

6.2 Case two

Un Balanced three-phase non-linear load is represented by three-phase uncontrolled diode bridge rectifier with pure resistive load at its DC bus. Fig. 10 shows the transient responses of distribution system with DSTATCOM for supply voltages (vsabc), supply currents (isabc), load currents (ila, ilb and ilc), DSTATCOM currents (ica, icb and icc) along with DC link voltage (Vdc) and its reference value (Vdcr) at rectifier nonlinear load.

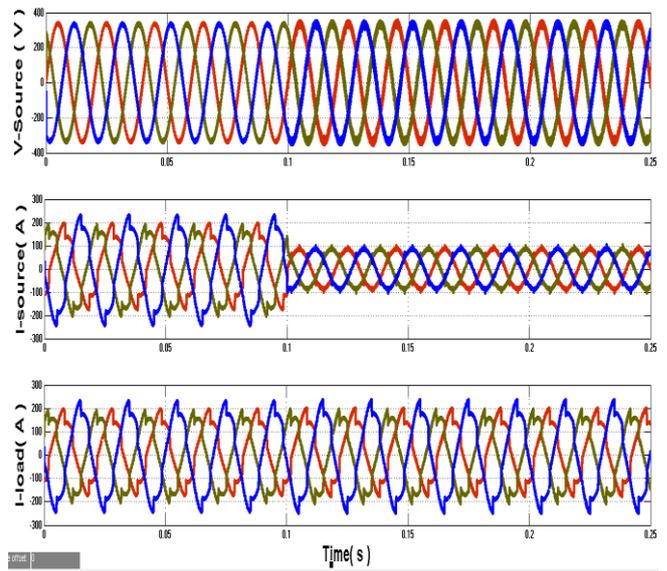


Figure. 10 Simulation results Non- Linear Unbalanced Load
 (a) source voltage (b) source current (c) load current

Fig.10 shows the unbalanced non linear load case. From the figure it is clear that even though load is unbalanced source currents are balanced and sinusoidal.

7.3 Case three

Performance of DSTATCOM connected to a weak supply system is shown in Fig.11 for power factor correction and load balancing. This figure shows variation of performance variables such as supply voltages (vsa, vsb and vsc), terminal voltages at PCC (vta, vtb and vtc), supply currents (isa, isb and isc), load currents (ila, ilb and ilc), DSTATCOM currents (ica, icb and icc) and DC link voltage (Vdc).

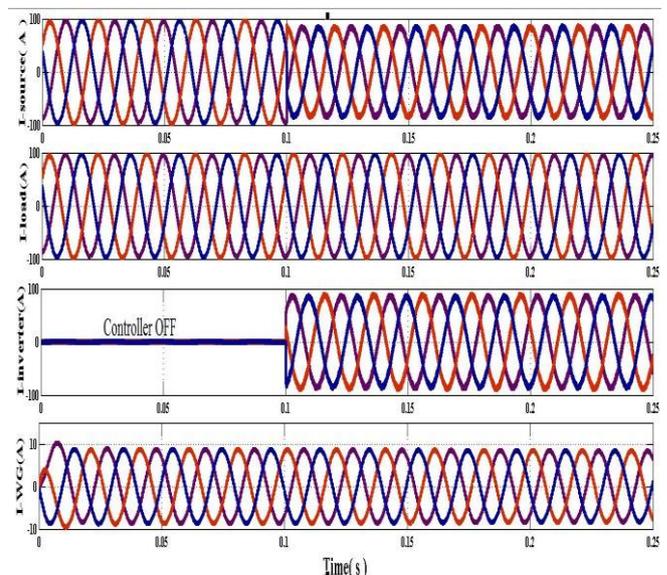


Figure. 11 Simulation results for Balanced Linear Load
 (a) Source current. (b) Load current. (c) Inverter injected current.
 (d) wind generator (induction generator) current.

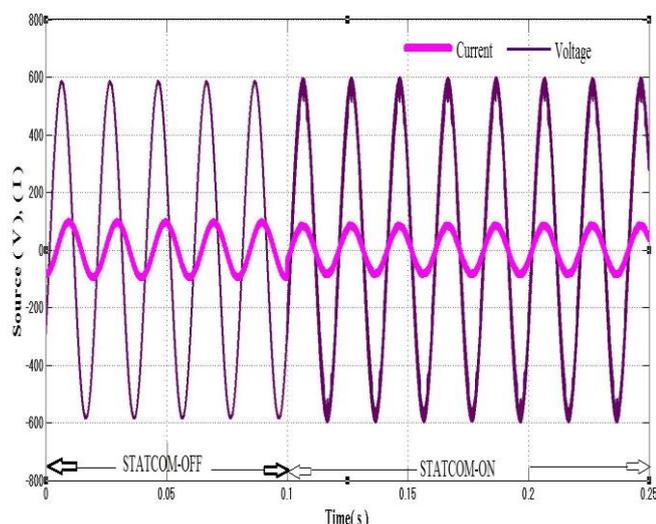


Figure. 12 Simulation results power factor for linear Load

Fig. 12 shows the power factor it is clear from the figure after compensation power factor is unity.

7.4 Case four

Un Balanced three-phase linear load is represented by three-phase uncontrolled diode bridge rectifier with pure resistive load at its DC bus. Fig. 10 shows the transient responses of distribution system with DSTATCOM for supply voltages (v_{sabc}), supply currents (i_{sabc}), load currents (i_{la} , i_{lb} and i_{lc}), inverter current (i_{ina} , i_{inb} , i_{inc}) DSTATCOM currents (i_{ca} , i_{cb} and i_{cc}) along with DC link voltage (V_{dc}) and its reference value (V_{dcr}) at rectifier linear load.

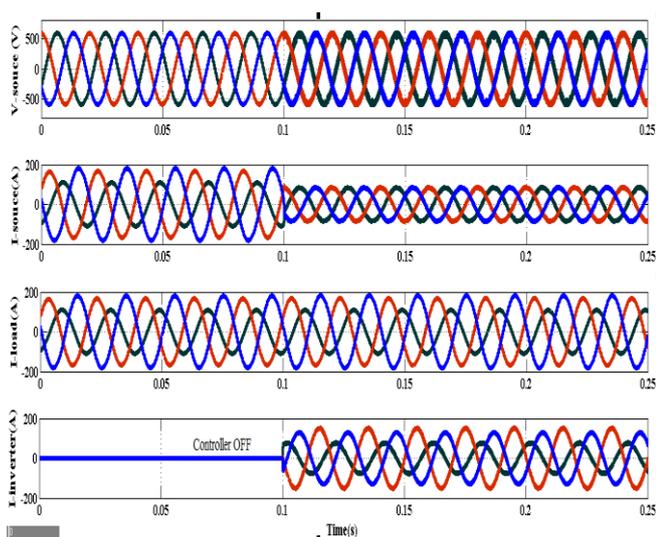


Figure. 13 Simulation results for unbalanced linear Load

VII. Conclusion

DSTATCOM system is an efficient mean for mitigation of PQ disturbances introduced to the grid by DERs. DSTATCOM compensator is a flexible device which can operate in current control mode for compensating voltage variation, unbalance and reactive power and in voltage control mode as a voltage stabilizer. The latter feature enables its application for compensation of dips coming from the supplying network. The simulation results show

that the performance of DSTATCOM system has been found to be satisfactory for improving the power quality at the consumer premises. DSTATCOM control algorithm is flexible and it has been observed to be capable of correcting power factor to unity, eliminate harmonics in supply currents and provide load balancing. It is also able to regulate voltage at PCC. The control algorithm of DSTATCOM has an inherent property to provide a self-supporting DC bus of DSTATCOM. It has been found that the DSTATCOM system reduces THD in the supply currents for non-linear loads. Rectifier-based non-linear loads generated harmonics are eliminated by DSTATCOM. When single-phase rectifier loads are connected, DSTATCOM currents balance these unbalanced load currents.

References

- [1] A.E. Hammad, Comparing the Voltage source capability of Present and future Var Compensation Techniques in Transmission System, IEEE Trans, on Power Delivery . volume 1. No.1 Jan 1995.
- [2] G.Yalienkaya, M.H.J Bollen, P.A. Crossley, "Characterization of Voltage Sags in Industrial Distribution System", IEEE transactions on industry applications, volume 34, No. 4, July/August, PP.682-688, 1999.
- [3] Haque, M.H., "Compensation Of Distribution Systems Voltage sags by DVR and D-STATCOM", Power Tech Proceedings, 2001 IEEE Porto, Volume 1, PP.10-13, September 2001.
- [4] Anaya-Lara O, Acha E., "Modeling and Analysis Of Custom Power Systems by PSCAD/EMTDC", IEEE Transactions on Power Delivery, Volume 17, Issue: 2002, Pages: 266-272.
- [5] Bollen, M.H.J., "Voltage sags in Three Phase Systems", Power Engineering Review, IEEE, Volume 21, Issue :9, September 2001, PP: 11-15.
- [6] M.Madriral, E.Acha., "Modelling OF Custom Power Equipment Using Harmonics Domain Techniques", IEEE 2000.
- [7] R.Meinski, R.Pawelek and I.Wasiak, "Shunt Compensation For Power Quality Improvement Using a STATCOM controller Modelling and Simulation", IEEE Proce, Volume 151, No. 2, March 2004.
- [8] J.Nastran , R. Cajhen, M. Seliger, and P.Jereb, "Active Power Filters for Nonlinear AC loads, IEEE Trans.on Power Electronics Volume 9, No.1, PP: 92-96, Jan 2004.
- [9] L.A.Moran, J.W. Dixon , and R.Wallace, A Three Phase Active Power Filter with fixed Switching Frequency For Reactive Power and Current Harmonics Compensation, IEEE Trans. On Industrial Electronics. Volume 42, PP:402-8, August 1995.
- [10] L.T. Moran ,P.D Ziogas, and G.Joos , Analysis and Design Of Three Phase Current source solid State Var Compensator, IEEE Trans, on Indutry Applications. Volume 25, No.2, 1989, PP:356-65.

Soret and Dufour Effects on MHD Free Convection Heat and Mass Transfer Flow over a Stretching Vertical Plate with Suction and Heat Source/Sink

M. J. Subhakar¹, K. Gangadhar²

¹ HOD of Mathematics, Noble College, Machilipatnam, A.P. 521001.India.

² Assistant Professor, Dept of mathematics, ANUOC, Ongole-523001, A.P, India.

ABSTRACT: The present paper is to investigate the combined effect of the free convective heat and mass transfer on the unsteady two-dimensional boundary layer flow over a stretching vertical plate in the presence of heat generation/absorption, and Soret and Dufour effects. The flow is subject to magnetic field normal to the plate. The governing nonlinear partial differential equations have been reduced to the coupled nonlinear ordinary differential equations by the similarity transformations. The resulting equations are solved numerically by using Runge-Kutta fourth order method along with shooting technique. The velocity and temperature distributions are discussed numerically and presented through graphs. The numerical values of skin-friction coefficient and Nusselt number at the plate are derived, discussed numerically for various values of physical parameters and presented through Tables. The numerical results are benchmarked with the earlier study by Magyari and Keller [24], and Srinivasachari and Ram Reddy [23] and found to be in excellent agreement.

Keywords: Free Convection, heat and mass transfer, heat generation/absorption, MHD, Suction, Stretching vertical plate.

I. INTRODUCTION

The effect of free convection on the accelerated flow of a viscous incompressible fluid past an infinite vertical plate with suction has many important technological applications in the astrophysical, geophysical and engineering problems. The heating of rooms and buildings by the use of radiators is a familiar example of heat transfer by free convection. Heat losses from hot pipes, ovens etc. Heat and mass transfer play an important role in manufacturing industries for the design of fins, steel rolling, nuclear power plants, gas turbines and various propulsion devices for aircraft, combustion and furnace design, materials processing, energy utilization, temperature measurements.

The unsteady heat transfer problems over a stretching surface, which is stretched with a velocity that depends on time, are considered by Andersson et al.[1] who studied the heat transfer over an unsteady stretching surface. Ishak et al.[2] have studied the heat transfer over an unsteady stretching vertical surface. Ishak et al.[3] have also investigated the unsteady laminar boundary layer over a continuously stretching permeable surface. Rajesh [4] has studied radiation effects on MHD free convection flow near a vertical plate with ramped wall temperature. El-Arabawy [5] studied the effect of suction/injection on a micropolar fluid past a continuously moving plate in the presence of radiation. Loganathan[6] studied the effects of thermal conductivity on unsteady MHD free convective flow over a semi infinite vertical plate.

In recent years, considerable attention has been devoted to study the MHD flows and heat transfer because of the applications in engineering, agriculture, petroleum industries, geophysics, and astrophysics. Kinyanjui et al. [7] studied the transient free convection heat over an impulsively started vertical plate in the presence of thermal radiation. Al-Odat and Al-Azab [8] numerically studied the transient MHD free convective

heat and mass transfer over a moving vertical surface in the presence of a homogeneous chemical reaction of first order. Palani and Srikanth [9] studied the MHD flow of an

electrically conducting fluid over a semi-infinite vertical plate under the influence of the transversely applied magnetic field. Makinde[10] investigated the MHD boundary layer flow with the heat and mass transfer over a moving vertical plate in the presence of magnetic field and convective heat exchange at the surface.

The heat source/sink effects in thermal convection are significant where there may exist high temperature differences between the surface (e.g. space craft body) and the ambient fluid. Heat generation is also important in the context of exothermic or endothermic chemical reaction. Tania et al [11] has investigated the Effects of radiation, heat generation and viscous dissipation on MHD free convection flow along a stretching sheet. Furthermore, Moalem [12] studied the effect of temperature dependent heat sources taking place in electrically heating on the heat transfer within a porous medium. Vajravelu and Nayfeh [13] reported on the hydro magnetic convection at a cone and a wedge in the presence of temperature dependent heat generation or absorption effects. Moreover, Chamkha [14] studied the effect of heat generation or absorption on hydro magnetic three-dimensional free convection flow over a vertical stretching surface. Radiation and mass transfer effects on MHD free convection fluid flow embedded in a porous medium with heat generation/absorption was studied by Shankar et al [15]

The Dufour and Soret effects were neglected in many reported research studies, since they are of a smaller order of magnitude than the effects described by Fourier's and Fick's laws. When the heat and mass transfer occurs simultaneously in a moving fluid, the relations between the fluxes and the driving potentials are more complicated. It was found that an energy flux can be generated not only by the temperature gradients but also by the composition gradients. The mass transfer caused by the temperature gradient is called the Soret effect, while the heat transfer caused by the concentration gradient is called the Dufour effect. However, such effects become crucial when the density difference exists in the flow regimes. The Soret effect, for instance, has been utilized for isotope separation.

In a mixture between gases with very light molecular weight (He, H₂) and medium molecular weight (N₂, air), the Dufour effect was found to be of considerable magnitude such that it cannot be neglected (Eckert and Drak [16]). The Dufour and Soret effects were studied by many researchers. Kafoussias and Williams [17] studied the thermo-diffusion and diffusion-thermo effects on the mixed free-forced convective and mass transfer of steady laminar boundary layer flow over a vertical flat. Mohamad [18] studied the Soret effect on the unsteady magnetohydrodynamics (MHD) free convection heat and mass transfer flow past a semi-infinite vertical plate in a Darcy porous medium in the presence of chemical reaction and heat generation. Afify[19] carried out a numerical analysis to study the free convective heat and mass transfer of an incompressible electrically conducting fluid over a stretching sheet in the presence of suction and injection with the Soret and Dufour effects. Shyam et al.[20] examined the Soret and Dufour effects on the MHD natural convection over a vertical surface embedded in a Darcy porous medium in the presence of thermal radiation. Numerical investigation of Dufour and Soret effects on unsteady MHD natural convection flow past vertical plate embedded in non-Darcy porous medium was investigated by Al-Odat and Al-Ghamdi [21]. Ali-Chamkha and Mansour[22] examined the effect of chemical reaction, thermal radiation, and heat generation or absorption on the unsteady MHD free convective heat and mass transfer along an infinite vertical plate. Soret and Dufour Effects on Mixed Convection from an Exponentially Stretching Surface was discussed by Srinivasacharya and RamReddy [23]

The object of the present chapter is to analyze the combined effect of the free convective heat and mass transfer on the unsteady two-dimensional boundary layer flow over a stretching vertical plate by taking heat generation/absorption, and soret and dufour effects into account. The governing boundary layer equations have been transformed to a two-point boundary value problem in similarity variables and the resultant problem is solved numerically using the Runge-Kutta method with shooting technique. The effects of various governing parameters on the fluid velocity, temperature, concentration, skin-friction coefficient, Nusselt number and Sherwood number are shown in figures and tables and analyzed in detail.

II. MATHEMATICAL ANALYSIS

An unsteady two-dimensional laminar boundary layer flow of an incompressible fluid over a stretching vertical plate is considered. The fluid is assumed to be Newtonian, electrically conducting and its property variations due to temperature and chemical species concentration are limited to fluid density. The density variation and the effects of the buoyancy are taken into account in the momentum equation (Boussinesq's approximation). The x-axis is taken along the plate in the upward direction and the y-axis is taken normal to it. A uniform magnetic field is applied in the direction perpendicular to the plate. The transverse applied magnetic field and magnetic Reynolds number are assumed to be very small, so that the induced magnetic field and Hall effects are negligible. Now, under the above assumptions, the governing boundary layer equations of the flow field are:

Continuity equation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad (2.1)$$

Momentum equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \nu \frac{\partial^2 u}{\partial y^2} - \frac{\sigma B_0^2}{\rho} u + g \beta_T (T - T_\infty) + g \beta_c (C - C_\infty) \quad (2.2)$$

Energy equation

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \alpha \frac{\partial^2 T}{\partial y^2} + \frac{D_m k_T}{c_s c_p} \frac{\partial^2 C}{\partial y^2} + q(T - T_\infty) \quad (2.3)$$

Species equation

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} = D_m \frac{\partial^2 C}{\partial y^2} + \frac{D_m k_T}{T_m} \frac{\partial^2 T}{\partial y^2} \quad (2.4)$$

The boundary conditions for the velocity, temperature and concentration fields are

$$u = U, v = V_w, T = T_w, C = C_w \text{ at } y = 0 \\ u \rightarrow 0, T \rightarrow T_\infty, C \rightarrow C_\infty \text{ as } y \rightarrow \infty \text{ for } t > 0 \quad (2.5)$$

where u , v , T and C are the fluid x-component of velocity, y-component of velocity, temperature and concentration respectively, ν is the fluid kinematics viscosity, ρ - the density, σ - the electric conductivity of the fluid, β_T and β_c - the coefficients of thermal and concentration expansions respectively, α - the thermal conductivity, C_∞ - the free stream concentration, B_0 - the magnetic induction, U - the free stream velocity, D_m - the mass diffusivity and g is the gravitational acceleration, T_w is the temperature of the hot fluid at the left surface of the plate, C_w is the species concentration at the plate surface, The mass concentration equation (2.1) is satisfied by the Cauchy-Riemann equations

$$u = \frac{\partial \psi}{\partial y}, v = -\frac{\partial \psi}{\partial x} \quad (2.6)$$

where $\psi(x, y)$ is the stream function.

To transform equations (2.2) - (2.4) into a set of ordinary differential equations, the following similarity transformations and dimensionless variables are introduced.

$$\eta = y \sqrt{\frac{c}{\nu(1-\lambda t)}}, \quad \psi = x \sqrt{\frac{c\nu}{1-\lambda t}} f(\eta) \\ T = T_\infty + T_w \left[\frac{cx}{2\nu(1-\lambda t)^2} \right] \theta(\eta) \\ C = C_\infty + C_w \left[\frac{cx}{2\nu(1-\lambda t)^2} \right] \phi(\eta) \\ M = \frac{\sigma B_0^2 (1-\lambda t)}{\rho c}, \quad Gr = \frac{g \beta_T T_w}{2c\nu}, \quad Gc = \frac{g \beta_c C_w}{2c\nu}, \\ D_f = \frac{D_m k_T C_w}{\nu c_s c_p T_w}, \quad Sr = \frac{D_m k_T T_w}{\nu T_m C_w}, \quad Q = \frac{cq}{1-\lambda t} \\ Pr = \frac{\nu}{\alpha}, \quad Sc = \frac{\nu}{D_m}, \quad A = \frac{\lambda}{c} \quad (2.7)$$

where η - similarity variable, f - dimensionless stream function, θ - dimensionless temperature, ϕ - dimensionless

concentration, M - the Magnetic field parameter, Gr - the thermal Grashof number, Gc - the solutal Grashof number, D_f - the Dufour number, Sr - the Soret number, Q - the heat generation/absorption parameter, Pr - the Prandtl number, Sc - the Schmidt number and A - Unsteadiness parameter.

In view of equations (2.6) and (2.7), Equations (2.2) to (2.4) transform into

$$f'''(\eta) + f(\eta)f''(\eta) - \frac{A}{2}\eta f'(\eta) + Gr\theta(\eta) + Gc\phi(\eta) - Mf'(\eta) = 0 \quad (2.8)$$

$$\theta''(\eta) + Pr \left[f(\eta)\theta'(\eta) - \frac{A}{2}\eta\theta'(\eta) - \theta f'(\eta) + Du\phi''(\eta) \right] = 0 \quad (2.9)$$

$$\phi''(\eta) + Sc \left[f(\eta)\phi'(\eta) - \frac{A}{2}\eta\phi'(\eta) - \phi(\eta)f'(\eta) + Sr\theta''(\eta) \right] = 0 \quad (2.10)$$

The corresponding boundary conditions are

$$\begin{aligned} f = fw, f' = \theta = \phi = 1 & \quad \text{at} & \quad y = 0 \\ f' = \theta = \phi = 0 & \quad \text{as} & \quad y \rightarrow \infty \end{aligned} \quad (2.11)$$

where the prime symbol represents the derivative with respect to η and fw is the suction parameter.

From the numerical computation, the local skin-friction coefficient, the local Nusselt number and Sherwood number which are respectively proportional to $f''(0)$, $-\theta'(0)$ and $-\phi'(0)$ are worked out and their numerical values are presented in a tabular form.

III. SOLUTION OF THE PROBLEM

The set of coupled non-linear governing boundary layer equations (2.8) - (2.10) together with the boundary conditions (2.11) are solved numerically by using Runge-Kutta fourth order technique along with shooting method. First of all, higher order non-linear differential Equations (2.8) - (2.10) are converted into simultaneous linear differential equations of first order and they are further transformed into initial value problem by applying the shooting technique (Jain *et al* [25]). The resultant initial value problem is solved by employing Runge-Kutta fourth order technique. The step size $\Delta\eta=0.05$ is used to obtain the numerical solution. From the process of numerical computation, the skin-friction coefficient, the Nusselt number and the Sherwood number, which are respectively proportional to $f''(0)$, $-\theta'(0)$ and $-\phi'(0)$, are also sorted out and their numerical values are presented in a tabular form.

IV. RESULTS AND DISCUSSION

The governing equations (2.9) - (2.11) subject to the boundary conditions (2.12) are integrated as described in section 3. Numerical results are reported in the Tables 1-2. The Prandtl number is taken to be $Pr = 0.71$ which corresponds to air, the value of Schmidt number (Sc) were chosen to be $Sc = 0.24, 0.62, 0.78, 2.62$, representing diffusing chemical species of most common interest in air like H_2 , H_2O , NH_3 and Propyl Benzene respectively.

The effects of various parameters on velocity profiles in the boundary layer are depicted in Figs. 1-10. In Fig. 1 the effect of increasing the magnetic field strength on the momentum boundary layer thickness is illustrated. It is now a well established fact that the magnetic field presents a damping effect on the velocity field by creating drag force that opposes the fluid motion, causing the velocity to decrease. Similar trend of slight decrease in the fluid velocity near the vertical plate is observed with an increase in suction parameter fw (see in Fig.2). Fig.3 illustrates the effect of the thermal Grashof number (Gr) on the velocity field. The thermal Grashof number signifies the relative effect of the thermal buoyancy force to the viscous hydrodynamic force. The flow is accelerated due to the enhancement in buoyancy force corresponding to an increase in the thermal Grashof number, i.e. free convection effects. It is noticed that the thermal Grashof number (Gr) influence the velocity field almost in the boundary layer when compared to far away from the plate. It is seen that as the thermal Grashof number (Gr) increases, the velocity field increases. The effect of mass (solutal) Grashof number (Gc) on the velocity is illustrated in Fig.4. The mass (solutal) Grashof number (Gc) defines the ratio of the species buoyancy force to the viscous hydrodynamic force. It is noticed that the velocity increases with increasing values of the solutal Grashof number.

Fig.5 illustrates the effect of the Schmidt number (Sc) on the velocity. The Schmidt number (Sc) embodies the ratio of the momentum diffusivity to the mass (species) diffusivity. It physically relates the relative thickness of the hydrodynamic boundary layer and mass-transfer (concentration) boundary layer. It is noticed that as Schmidt number (Sc) increases the velocity field decreases. Similar trend of slight decrease in the fluid velocity near the vertical plate is observed with an increase in Prandtl number Pr (see in Fig.6). Fig. 7 shows the variation of the velocity boundary-layer with the heat generation/absorption parameter (Q). It is noticed that the velocity boundary layer thickness increases with an increase in the generation/absorption parameter. Fig. 8 shows the variation of the velocity boundary-layer with the Dufour number (D_f). It is observed that the velocity boundary layer thickness increases with an increase in the Dufour number. Fig. 9 shows the variation of the velocity boundary-layer with the Soret number (Sr). It is found that the velocity boundary layer thickness increases with an increase in the Soret number. Fig. 10 shows the variation of the velocity boundary-layer with the Unsteadiness parameter (A). It is found that the velocity boundary layer thickness increases with an increase in the Unsteadiness parameter.

As per the boundary conditions of the flow field under consideration, the fluid temperature attains its maximum value at the plate surface and decreases exponentially to the free stream zero value away from the plate. This is observed in Figs. 11-20. The effect of the magnetic parameter M on the temperature is illustrated in Fig.11. It is observed that as the magnetic parameter M increases, the temperature increases. Fig.12 illustrates the effect of the suction parameter (fw) on the temperature. It is noticed that as suction parameter increases, the temperature decreases. From Figs. 13 and 14, it is observed that the thermal boundary layer thickness decreases with an increase in the thermal or Solutal Grashof numbers (Gr or

Gc). Fig. 15 illustrates the effect of Schmidt number (Sc) on the temperature. It is noticed that as the Schmidt number (Sc) increases, an increasing trend in the temperature field is noticed. Much of significant contribution of Schmidt number (Sc) is noticed as we move far away from the plate.

Fig.16 illustrates the effect of the Prandtl number (Pr) on the temperature. It is noticed that as the Prandtl number increases, the temperature decreases. Fig.17 illustrates the effect of the heat generation/absorption parameter (Q) on the temperature. It is noticed that as the heat generation/absorption parameter increases, the temperature increases. Fig. 18 shows the variation of the thermal boundary-layer with the Dufour number (D_f). It is noticed that the thermal boundary layer thickness increases with an increase in the Dufour number. Fig. 19 shows the variation of the thermal boundary-layer with the Soret number (Sr). It is observed that the thermal boundary layer thickness decreases with an increase in the Soret number. The effect of Unsteadiness parameter (A) on the temperature field is illustrated Fig.20. As the Unsteadiness parameter increases the thermal boundary layer is found to be increasing.

Figs. 21 - 28 depict chemical species concentration against span wise coordinate η for varying values of physical parameters in the boundary layer. The species concentration is highest at the plate surface and decrease to zero far away from the plate satisfying the boundary condition. The effect of magnetic parameter (M) on the concentration field is illustrated in Fig.21. As the magnetic parameter increases the concentration is found to be increasing. The effect of suction parameter (f_w) on the concentration field is illustrated Fig.22. As the suction parameter increases the concentration is found to be decreasing. However, as we move away from the boundary layer, the effect is not significant. The effect of buoyancy parameters (Gr , Gc) on the concentration field is illustrated Figs. 23 and 24. It is noticed that the concentration boundary layer thickness decreases with an increase in the thermal or Solutal Grashof numbers (Gr or Gc). Fig. 25 illustrates the effect of Schmidt number (Sc) on the concentration. It is noticed that as the Schmidt number increases, there is a decreasing trend in the concentration field.

The influence of the heat generation/absorption parameter (Q) on the concentration field is shown in Fig.26. It is noticed that the concentration decreases monotonically with the increase of the heat generation/absorption parameter. Fig. 27 shows the variation of the concentration boundary-layer with the Dufour number (D_f). It is observed that the concentration boundary layer thickness decreases with an increase in the Dufour number. Fig. 28 shows the variation of the concentration boundary-layer with the Soret number (Sr). It is found that the concentration boundary layer thickness increases with an increase in the Soret number.

In order to benchmark our numerical results, the present results for the skin -friction, Nusselt number and Sherwood number in the absence of M , A , f_w , Q , D_f , Sr and S are compared with those of Magyari and Keller [24] and Srinivasachary and Ram Reddy [23] and found them in excellent agreement as demonstrated in Table 1. From Table 2, it is observed that the local skin-friction coefficient, local heat and mass transfer rates at the plate

increases with an increase in the buoyancy forces. It was noticed that the local skin-friction coefficient, local heat and mass transfer rates at the plate decreases with an increase in the Magnetic parameter. As the Prandtl number increases, both the skin-friction and Sherwood number decrease, whereas the Nusselt number increases. It was found that the local heat and mass transfer rate at the plate increases, but Skin-friction coefficient decreases with an increase in the Schmidt number or suction parameter. It is observed that the local heat and mass transfer rate at the plate decreases, whereas Skin-friction coefficient increases with an increase in the Unsteadiness parameter. As the Dufour number or heat generation/absorption parameter increases, both the skin-friction and Sherwood number increase, whereas the Nusselt number decreases. The effect of the Soret number is to increase the Skin-friction and the Nusselt number, and to decrease the Sherwood number.

V. CONCLUSIONS

The present chapter analyzes the combined effect of the free convective heat and mass transfer on the unsteady two-dimensional boundary layer flow over a stretching vertical plate by taking heat generation/absorption, and soret and dufour effects into account. The governing equations are approximated to a system of non-linear ordinary differential equations by similarity transformation. Numerical calculations are carried out for various values of the dimensionless parameters of the problem. A comparison with previously published work is performed and excellent agreement between the results is found. The results are presented graphically and the conclusion is drawn that the flow field and other quantities of physical interest are significantly influenced by these parameters. It is noticed that as the heat generation/absorption parameter increases, the velocity and temperature increases, where as concentration decreases. The results for the prescribed skin friction, local heat and mass transfer rates at the plate are presented and discussed. It is found that the local skin-friction coefficient, local heat and mass transfer rates at the plate increase with an increase in the buoyancy forces. As the heat generation/absorption parameter or dufour number increases, both the skin-friction and Sherwood number increase, whereas the Nusselt number decreases. It is observed that the local skin-friction coefficient and local mass transfer rates at the plate decrease but Nusselt number decreases with an increase in the Prandtl number. The effect of the Soret number is to increase the local heat and mass transfer rates, and to decrease the skin-friction.

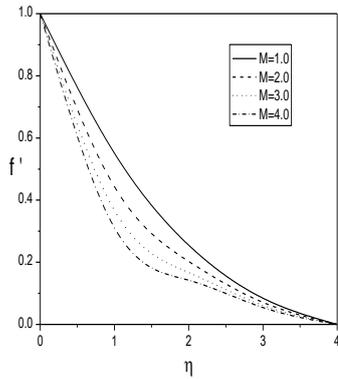


Fig.1: Variation of the velocity f' with M for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, Q=0.2, D_f=Sr=fw=0.1$.

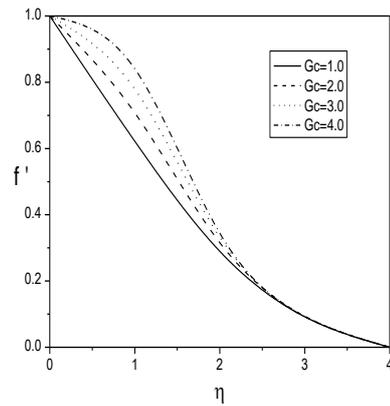


Fig.4: Variation of the velocity f' with Gc for $Pr=0.71, Sc=0.24, Gr=A=1, Q=0.2, M=0.5, D_f=Sr=fw=0.1$.

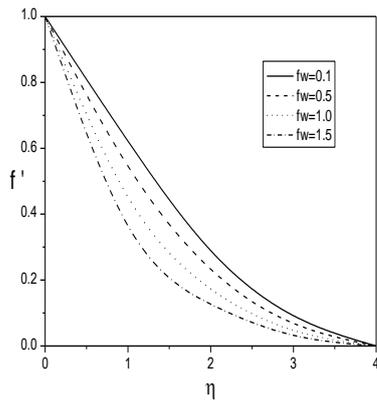


Fig.2: Variation of the velocity f' with fw for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, M=0.5, Q=0.2, D_f=Sr=0.1$.

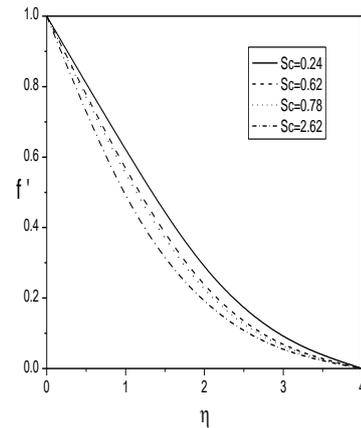


Fig.5: Variation of the velocity f' with Sc for $Pr=0.71, Gr=Gc=A=1, Q=0.2, M=0.5, D_f=Sr=fw=0.1$.

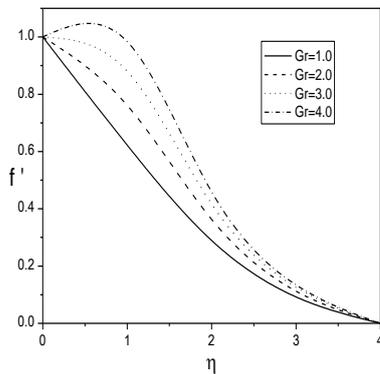


Fig.3: Variation of the velocity f' with Gr for $Pr=0.71, Sc=0.24, Gc=A=1, M=0.5, Q=0.2, D_f=Sr=fw=0.1$.

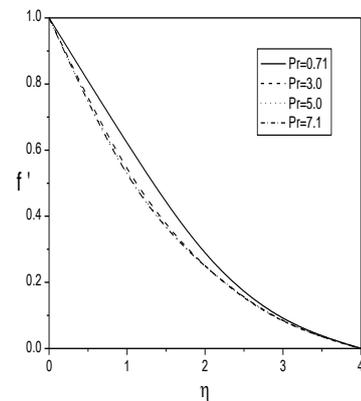


Fig.6: Variation of the velocity f' with Pr for $Sc=0.24, Gr=Gc=A=1, Q=0.2, M=0.5, D_f=Sr=fw=0.1$.

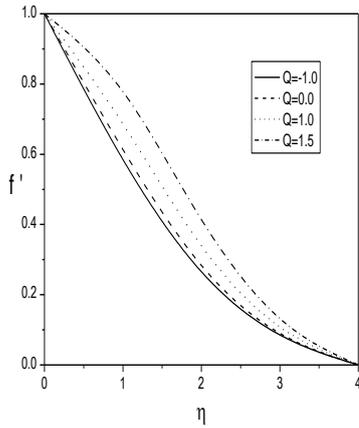


Fig.7: Variation of the velocity f' with Q for $Sc=0.24, Pr=0.71, Gr=Gc=A=1, M=0.5, D_f=Sr=f_w=0.1$.

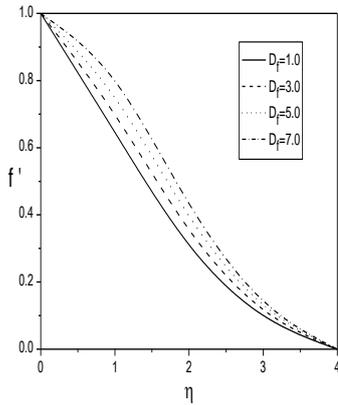


Fig.8: Variation of the velocity f' with D_f for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, M=0.5, Q=0.2, Sr=f_w=0.1$.

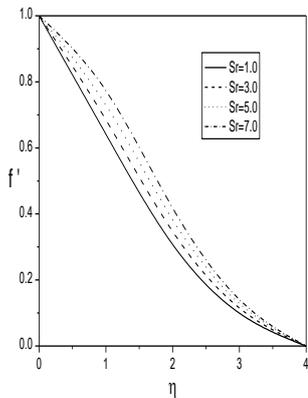


Fig.9: Variation of the velocity f' with Sr for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, M=0.5, Q=0.2, D_f=f_w=0.1$.

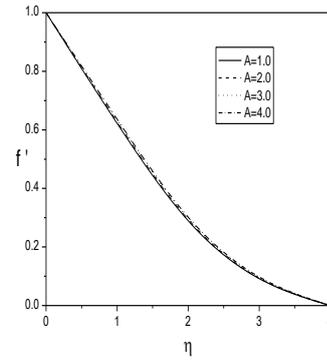


Fig.10: Variation of the velocity f' with Sr for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, M=0.5, Q=0.2, D_f=f_w=0.1$.

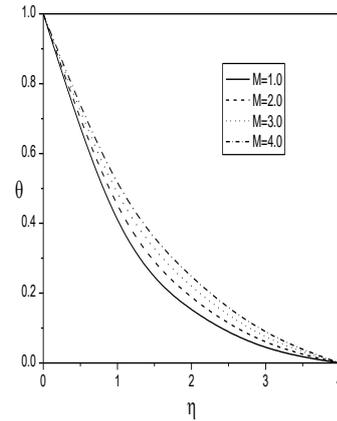


Fig.11: Variation of the temperature θ with M for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, Q=0.2, D_f=Sr=f_w=0.1$.

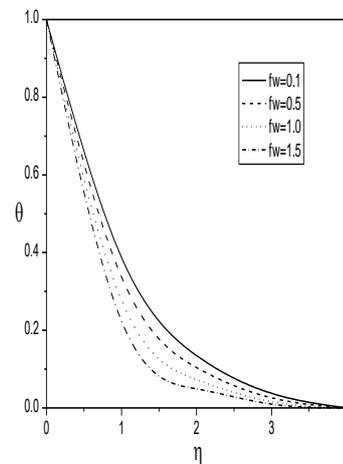


Fig.12: Variation of the temperature θ with f_w for $Pr=0.71, Sc=0.24, Gr=Gc=A=1, M=0.5, Q=0.2, D_f=Sr=0.1$.

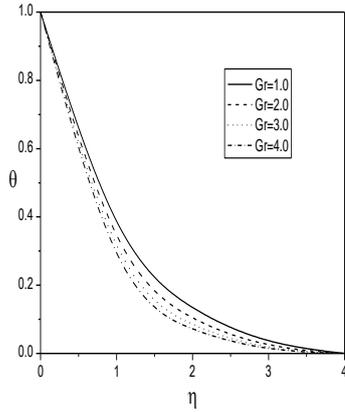


Fig.13: Variation of the temperature θ with Gr for $Pr=0.71$, $Sc=0.24$, $Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

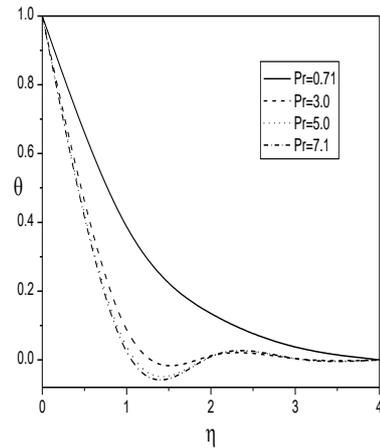


Fig.16: Variation of the temperature θ with Pr for $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

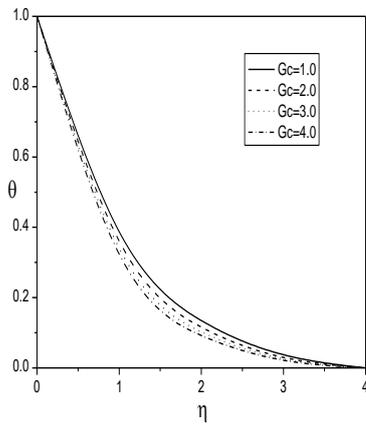


Fig.14: Variation of the temperature θ with Gc for $Pr=0.71$, $Sc=0.24$, $Gr=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

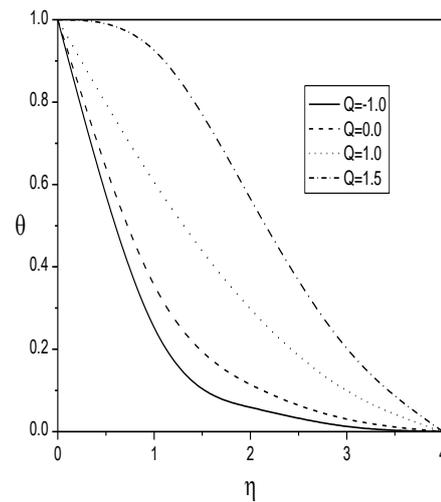


Fig.17: Variation of the temperature θ with Q for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $D_f=Sr=fw=0.1$.

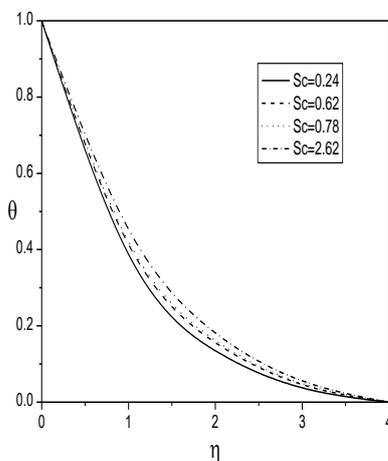


Fig.15: Variation of the temperature θ with Sc for $Pr=0.71$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

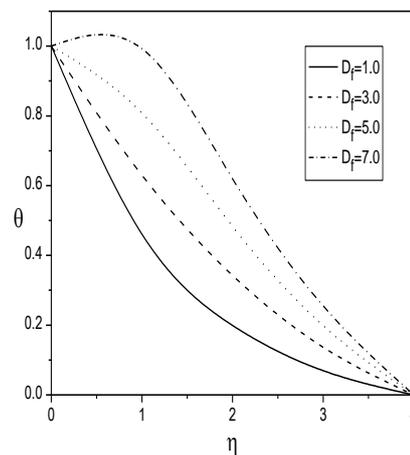


Fig.18: Variation of the temperature θ with D_f for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $Sr=fw=0.1$.

for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $Q=0.2$,
 $D_f=Sr=fw=0.1$.

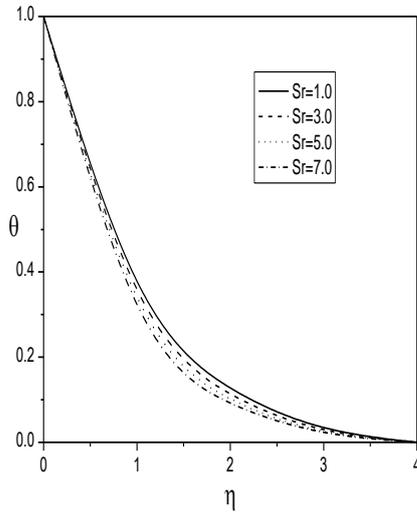


Fig.19: Variation of the temperature θ with Sr for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=fw=0.1$.

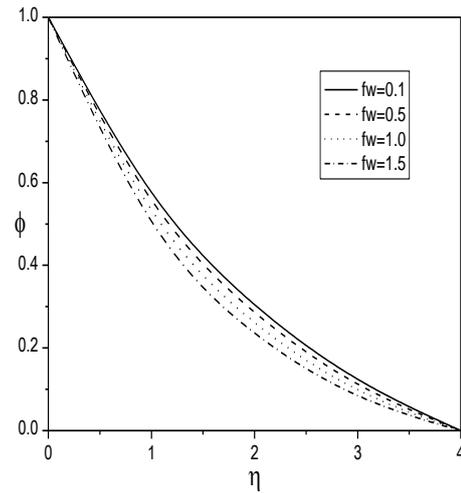


Fig.22: Variation of the concentration ϕ with fw for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=0.1$.

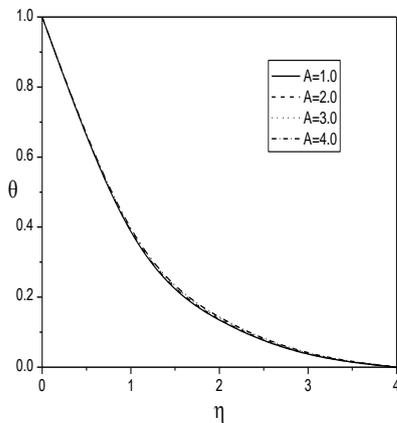


Fig.20: Variation of the temperature θ with A for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

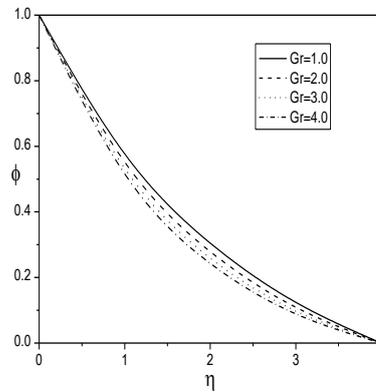


Fig.23: Variation of the concentration ϕ with Gr for $Pr=0.71$, $Sc=0.24$, $Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

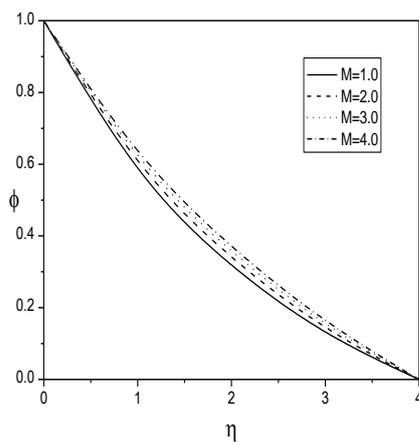


Fig.21: Variation of the concentration ϕ with M

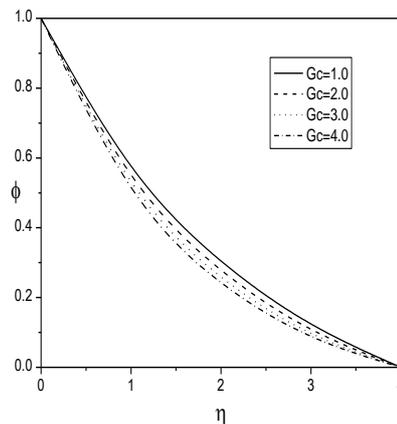


Fig.24: Variation of the concentration ϕ with Gc for $Pr=0.71$, $Sc=0.24$, $Gr=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

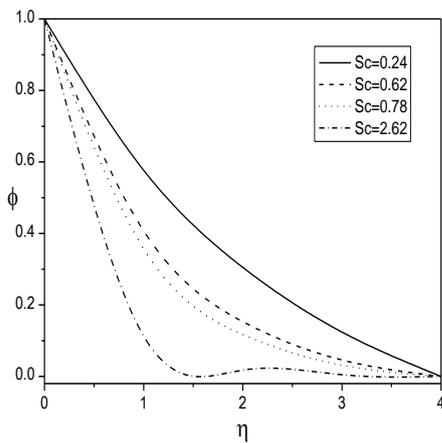


Fig.25: Variation of the concentration ϕ with Sc for $Pr=0.71$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=Sr=fw=0.1$.

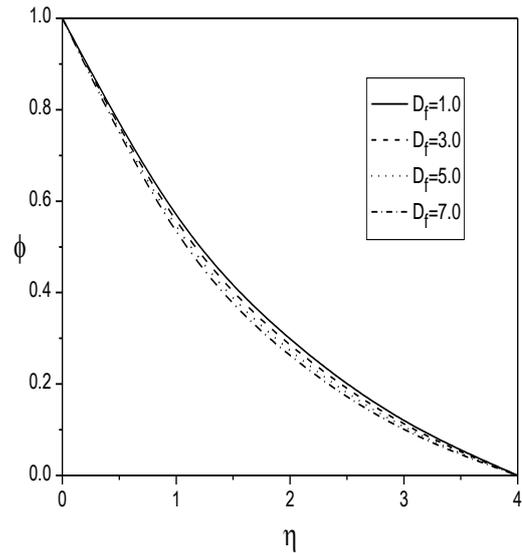


Fig.27: Variation of the concentration ϕ with D_f for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $Sr=fw=0.1$.

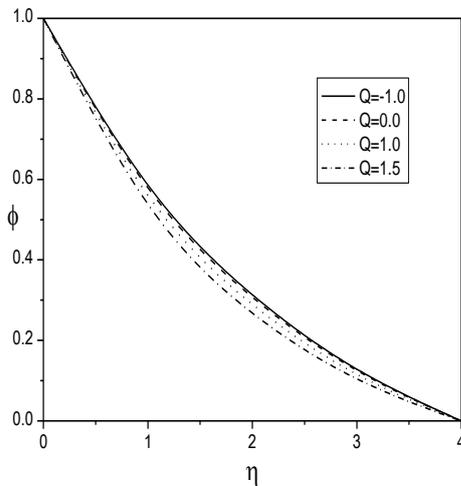


Fig.26: Variation of the concentration ϕ with Q for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $D_f=Sr=fw=0.1$.

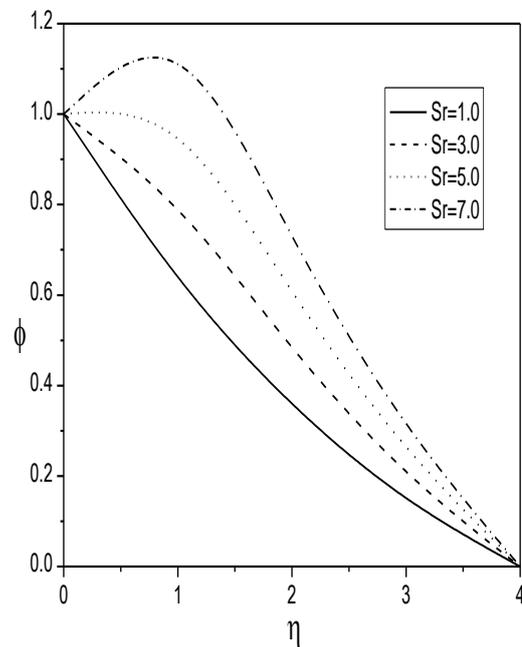


Fig.28: Variation of the concentration ϕ with Sr for $Pr=0.71$, $Sc=0.24$, $Gr=Gc=A=1$, $M=0.5$, $Q=0.2$, $D_f=fw=0.1$.

able 1 Numerical values of $\theta'(0)$ for $Gr=Gc=Sr=D_f=Q=fw=0$ and $Sc \rightarrow 0$. Comparison of the present results with that of Magyari and keller [24], and Srinivasachary and Ram Reddy [23].

Pr	Magyari and keller [24]	Srinivasachary and Ram Reddy [23]	Present Results
0.5	-0.59434	-0.59438	-0.667192
1.0	-0.95478	-0.95478	-1.00426
3.0	-1.86908	-1.86908	-1.92168
5.0	-2.50014	-2.50015	-2.55594

8.0	-3.24213	-3.24218	-3.30016
10.0	-3.66038	-3.66043	-3.71928

Table 2 Variation of $f''(0)$, $-\theta'(0)$ and $-\phi'(0)$ at the plate with Gr , Gc , M , Pr , Sc , Sr , D_f , A , Q , fw .

Gr	Gc	M	Pr	Sc	Sr	D_f	A	Q	fw	$f''(0)$	$-\theta'(0)$	$-\phi'(0)$
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.268957	0.882380	0.527214
2	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01	0.193829	0.949762	0.560834
1	2	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01	0.115906	1.002640	0.588621
1	3	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01	0.478782	0.949018	0.549084
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.504938	0.967315	0.567495
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01	-1.243580	0.844139	0.509469
1	1	0	1.01	0.24	0.01	0.01	1	0.02	0.01	-0.418908	0.723913	0.459550
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.342720	2.013790	0.483456
1	1	3	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.364602	2.706920	0.464043
1	1	5	3.0	0.24	0.01	0.01	1	0.02	0.01	-1.179407	0.387119	0.856410
1	1	0.5	5.0	0.24	0.01	0.01	1	0.02	0.01	-0.114296	0.822721	0.972591
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.160599	0.933253	0.114921
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.084164	0.968535	-0.200696
1	1	0.5	0.71	0.24	0.01	0.01	2	0.02	0.01	-0.260365	0.406943	0.555048
1	1	1	0.71	0.24	0.01	0.01	3	0.02	0.01	-0.251900	0.046733	0.574042
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.241963	0.877468	0.525047
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.172646	0.872552	0.522894
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.428342	0.7392	
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01	-0.684334		
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	0.5	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	1	0.71	0.24	0.01	0.01	1	0.02	0.01			
1	1	5	0.71	0.24								

		5	0.7 1	4	1	1		5	1		13	0.53437 2
		0.		0.2	0.	0.		1.	0.		0.4188	
		5	0.7 1	4	1	1		0	1		66	0.55196 4
		0.		0.2	0.	0.		0.	0.		1.0279	
		5	0.7 1	4	1	1		2	5		40	0.56210 6
		0.		0.2	0.	0.		0.	1.		1.2347	
		5		4	1	1		2	0		70	0.60945 1

REFERENCES

- Anderson, H.T., Aarseth, J.B. and Dandapat, B.S. (2000), Heat trans-fer in a liquid film on an unsteady stretching surface, *Int. J. Heat Transf.* Vol. 43, pp. 69-74.
- Ishak, A., Nazar, R., and Pop, I., (2007), "Magneto-hydrodynamic Stagnation Point Flow. Towards a Stretching Vertical Sheet in a Micropolar Fluid", *Magneto-hydrodynamics*, Vol. 43, pp. 83 – 97.
- Ishak, A., Nazar, R., and Pop, I. (2006), "Unsteady Mixed Convection Boundary Layer Flow Due to a Stretching Vertical Surface", *the Arabian Journal for Science and Engineering*, Vol. 31, pp. 165 – 182.
- Rajesh, V. (2010), MHD effects on free convection and mass transform flow through a porous medium with variable temperature, *Int. J. of Appl. Math and Mech.*, Vol. 6, No. 14, pp.1 – 16.
- El-Arabawy, H. A. M. (2003), Effect of Suction/Injection on a Micropolar Fluid Past a Continuously Moving Plate in the Presence of Radiation, *Int. J. Heat Mass Trans.*, vol.46, pp.1471-1477.
- Loganathan, P. (2010), Effects of Thermal Conductivity on Unsteady Mhd Free Convective Flow Over A Semi Infinite Vertical Plate, *International Journal of Engineering Science and Technology*, Vol. 2, No. 11, pp. 6257-6268.
- Kinyanjui, M., Wanza, J. K. K., and Uppal, S. M. (2001), Magneto-hydrodynamic free convection heat and mass transfer of a heat generating fluid past an impulsively started infinite vertical porous plate with Hall current and radiation absorption, *Energy Conversion and Management*, Vol. 42, No. 2, pp. 917–931.
- Al-Odat, M. Q. and Al-Azab, T. A. (2007), Influence of chemical reaction on transient MHD free convection over a moving vertical plate, *Emirates Journal for Engineering Research*, Vol. 12, No. 3, pp. 15–21.
- Palani, G. and Srikanth, U. (2009), MHD flow past a semi-infinite vertical plate with mass transfer. *Nonlinear Analysis, Modelling and Control*, Vol. 14, No. 3, pp. 345–356.
- Makinde, O. D. (2010), On MHD heat and mass transfer over a moving vertical plate with a convective surface boundary condition, *The Canadian Journal of Chemical Engineering*, Vol. 88, No. 6, pp. 983–990.
- Tania S. Khaleque and Samad M.A. (2010), Effects of radiation, heat generation and viscous dissipation on MHD free convection flow along a stretching sheet, *Research J of Appl. Sci., Eng. and Tech.*, Vol. 2, No. 4, pp. 368-377.
- Moalem .D. (1976), Steady state heat transfer with porous medium with temperature dependent heat generation, *Int. J.Heat and Mass Transfer*, Vol. 19, No. 529–537.
- Vajrevelu. K., J. Nayfeh, (1992), Hydro magnetic convection at a cone and a wedge, *Int. Comm. Heat Mass Transfer*, Vol. 19, pp. 701–710.
- Chamkha A. J. (1999), Hydro magnetic three-dimensional free convection on a vertical stretching surface with heat generation or absorption, *Int. J. Heat and Fluid Flow*, Vol. 20, pp. 84–92.
- Shankar, B., Prabhakar Reddy, B. And Ananda Rao, J. (2010), Radiation and mass transfer effects on MHD free convection fluid flow embedded in a porous medium with heat generation/absorption, *Indian Journal of pure and applied physics*, Vol. 48, pp.157-165.
- Eckert, E. R. G. and Drak, R. M. (1972), *Analysis of Heat and Mass Transfer*, McGraw-Hill, New York.
- Kafoussias, N. G. and Williams, E. W. (1995), Thermal-diffusion and diffusion-thermo effects on mixed free-forced convective and mass transfer boundary layer flow with temperature dependent viscosity. *International Journal of Engineering Science*, Vol. 33, No. 9, pp. 1369–1384.
- Mohamad, R. A. (2009), Double-diffusive convection-radiation interaction on unsteady MHD flow over a vertical moving porous plate with heat generation and Soret effect, *Applied Mathematical Sciences*, Vol. 3, No. 13, pp. 629–651.
- Afify, A. A. (2009), Similarity solution in MHD: effects of thermal-diffusion and diffusion-thermo on free convective heat and mass transfer over a stretching surface considering suction or injection, *Communications in Nonlinear Science and Numerical Simulation*, Vol. 14, No. 5, pp. 2202–2214.
- Shyam, S. T., Rajeev, M., Rohit, K. G., and Aiyub, K. (2010), MHD free convection-radiation interaction along a vertical surface embedded in Darcian porous medium in presence of Soret and Dufour's effects, *Thermal Science*, Vol. 14, No. 1, pp.137–145.
- Al-Odat, M.Q. and Al-Ghamdi, A. (2012), Dufour and Soret effects on unsteady MHD natural convection flow past vertical plate embedded in non-Darcy porous medium, *Appl. Math. Mech. -Engl. Ed.*, Vol.33, No.2, pp.195–210.
- Ali-Chamkha, M. A. and Mansour, A. A. (2011), Unsteady MHD free convective heat and mass transfer from a vertical porous plate with Hall current, thermal radiation and chemical reaction effects, *International Journal for Numerical Methods in Fluids*, Vol.65, No.4, pp.432–447.
- Srinivasacharya, D. and RamReddy, Ch. (2011), Soret and Dufour Effects on Mixed Convection from an Exponentially Stretching Surface, *International Journal of Nonlinear Science*, Vol.12, No.1, pp.60-68.
- Magyari, E. and Keller, B (1999), Heat and mass transfer in the boundary layers on an exponentially stretching continuous surface, *J. Phys. D: Appl. Phys.*, Vol.32, and No.577-585.
- Jain, M.K., Iyengar, S.R.K., and Jain, R.K. (1985), *Numerical Methods for Scientific and Engineering Computation*, Wiley Eastern Ltd., New Delhi, India.

FPGA Implementation of the Block RLS Algorithm

S.VijayaLakshmi¹, K.Raghuram²

¹. M.Tech Scholar, E.C.E, Pragati Engineering College, Surampalem, JNTUK, India

². Associate professor, E.C.E, JNTUK

ABSTRACT: Recursive Least Square (RLS) algorithm which recursively find the filter coefficients that minimises a weighted linear least square cost function. Recent digital transmission systems impose the application of channel equalizers with short training time and high tracking rate. These requirements turn our attention to adaptive algorithms, which converge rapidly. The demand for fast convergence and less MSE level cannot be met by conventional adaptive filtering algorithms such as LMS. The best choice is the block recursive least squares (RLS) algorithm. Block Recursive Least Squares algorithms are known to exhibit better performances. In this paper we present a simple architecture for the implementation of a variant of Block RLS algorithm where the weight updation and error calculation are both calculated block wise. The performance of the Simplified BRLS and LMS algorithms are compared in MATLAB simulations and the hardware outputs from the FPGA are verified with the simulations.

Keywords: RLS, ALTERA, Modelsim, Quartus, Matlab.

I. INTRODUCTION

In spite of the computation efficiency of the algorithm, additional simplifications may be necessary in some applications. There are several approaches that may be used to reduce the computational requirements of the RLS algorithm. One of these is the block RLS algorithm [2]. Block digital filtering has been extensively discussed by Burrus, Mitra and others. The technique involves calculation of a block of finite set of output values from a block of input values. Block implementations of digital filters allow efficient use of parallel processors which can result in speed gains [1]

RLS is one of the adaptive filtering algorithms derived from steepest descent algorithm used in wide variety of applications including System Identification, Channel Equalization, Adaptive antennas, and Noise cancellation. There are several variants of RLS which are selected depending on the system constraints like speed, hardware efficiency, accuracy of measurements etc. Block RLS is one of the variants in which the weights are updated once per every block data instead of updating on every clock cycle of input data. In block RLS, though the updation is done at block level the error values are calculated every clock cycle and are averaged over the block of the data. In this paper we further simplified these calculations and performed the error calculation also, once per block of data, making it a block implementation in the true sense and making it affable for hardware implementation.

The trade off for this approximation is the increased delay for convergence of weights and increase in oscillations of the error square plot. But with this approximation the algorithm could be easily implemented on FPGA using minimal hardware. The rest of the paper is

Organized as follows. Section III describes the architecture of implementation. Section IV briefs on simulation results. Section V shows the hardware utilization summary; Section VI contains the conclusion of the paper.

II. ADAPTIVE SYSTEM IDENTIFICATION

The adaptive system identification is primarily responsible for determining a discrete estimation of the transfer function for an unknown digital or analog system. The same input $x(n)$ is applied to both the adaptive filter and the unknown system from which the outputs are compared (see figure 1). The output of the adaptive filter $y(n)$ is subtracted from the output of the unknown system resulting in a desired signal $d(n)$. The resulting difference is an error signal $e(n)$ used to manipulate the filter coefficients of the adaptive system trending towards an error signal of zero. After a number of iterations of this process are performed, and if the system is designed correctly, the adaptive filter's transfer function will converge to, or near to, the unknown system's transfer function. For this configuration, the error signal does not have to go to zero, although convergence to zero is the ideal situation, to closely approximate the given system. There will, however, be a difference between adaptive filter transfer function and the unknown system transfer function if the error is nonzero and the magnitude of that difference will be directly related to the magnitude of the error signal.

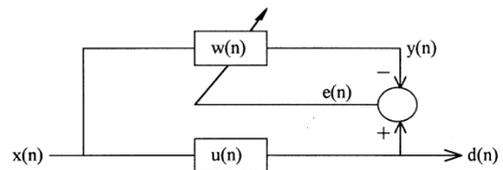


Figure1. Adaptive filter.

A. Adaptive Filters in System Identification Mode:

Since there are many applications for adaptive filters such as those just described, we often alter the structure of the filter used to suit the application. In this project we are interested in the use of adaptive filters for the purposes of system identification in which case a block diagram of the filter will be of the form shown in Figure 2. Notice that the error signal in this case is the difference between the filter output and the desired response, the output of the unknown system.

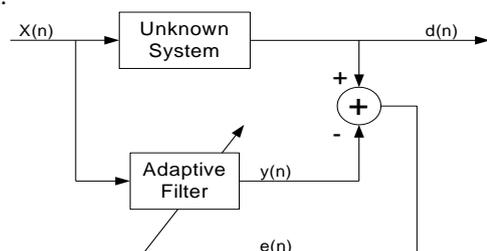


Figure2. An adaptive filter in system identification mode

B. Convolution:

In this sub-section, we will define the above term for the purposes of this project. Let us first define Correlation, in essence the Cross Correlation of two waveforms is a method for comparing the waveforms. Consider two waveforms, both sampled at the same rate. The sum of the products of the corresponding pairs of points is represented as a measure of the correlation of the two waveforms. Convolution can now be introduced as the cross correlation of two signals with one of them reversed [4]. Equation 1.0 defines the process for continuous time signals.

$$y(t) = x(t) * h(t) \tag{1.0}$$

where * denotes convolution, when $x(t)$ and $h(t)$ are two finite digital signals this Equation can be represented as

$$y(n) = \sum_{k=0}^{N-i} h(k) - x(n - k) \tag{1.1}$$

Notice that this is the same as Equation 1.1, which describes the operation of digital filters. In this case one of the sequences to be convolved is the impulse Response of the digital filter. Therefore it can be recognized that digital filtering is an application of convolution. Convolution can be viewed as a description of how the input to a given system reacts with the system's impulse response to produce an output. Convolution in the time domain is very computationally intensive and the fact that the convolution of two sequences can be calculated more efficiently if the signals are transformed from the time domain to the frequency domain is exploited in following chapters in an attempt to reduce the computation necessary for digital filters.

C. Recursive and non Recursive Filters:

Non-recursive filters are filters where the output depends only on current and previous input samples as depicted in Figure 3. An important property of non-recursive filters is they will always be stable. FIR filters are in general non-recursive, which in turn means that they are also always stable [2].

A recursive filter on the other hand is a filter whose output samples may depend on previous output samples as well as the current and previous input samples.

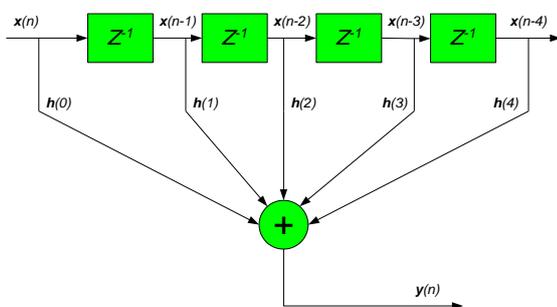


Figure 3. A Non-Recursive filter structure

Having introduced the concept of digital filters and the fact that the basis for digital filtering is time domain convolution, in particular stable FIR filters.

D. Digital filtering (Echo cancellation):

The hybrid is basically a bridge circuit with three ports and it is here that impedance mismatches can occur if the bridge is not perfectly balanced [1]. The adaptive filter is in essence the echo canceller. This filter operates in system identification mode with the echo path representing the unknown system the filter is to identify. Notice the inputs to the adaptive filter, speech from the transmitter constitutes the desired response and Equation 1.2 represents the error signal.

$$e(n) = s(n) - y(n) \tag{1.2}$$

where $y(n)$ is the output of the adaptive filter and $s(n)$ is composed of the other speaker's speech and the echo of the speaker's own speech. For satisfactory operation of the echo canceller, the length of the filter's impulse response must be greater than the length of the longest echo path.

Having introduced the project in the previous chapter it is time to introduce digital filtering. Adaptive filters are mainly implemented with digital filters, digital filtering therefore has a significant role to play in this project.

1) Introduction to Digital Filters:

The system or network in the case of digital filters are mathematical algorithms, these algorithms operate on digital signals and attempt to modify these signals in the same way analogue filters modify analogue signals. Equation 1.3 defines the operation of linear digital filters [2].

$$y(n) = \sum_{k=0}^{N-1} h(k) - x(n - k) \tag{1.3}$$

where $h(k)$, $k = 0, 1 \dots N-1$ are the filter coefficients, $x(n)$ is the filter input and $y(n)$ is the filter output. We note at this point that the above Equation represents the convolution of the input signal $x(n)$ with the filter's impulse response $h(k)$ to give the filter output $y(n)$. There will be further discussion of convolution in Section 2.2.

- Automatic updating of frequency Response if a programmable processor is used, which means that adaptive filtering can be implemented more easily.
- The filter output and input can be stored for further use.
- Some characteristics are not possible with analogue filters.
- Can easily take advantage of advances in VLSI technology.
- Performance is not affected by temperature variations or by slight differences in components making it possible to repeat the characteristics from one filter to the next.
- More than one signal can be filtered at a time.
- Precision is only limited by word length.
- It is possible to use digital filters at lower frequencies, which are often found in biomedical applications.

2) Disadvantages of Digital Filtering:

- The speed at which the filter operates may be limited by the speed of the processor or by the number of arithmetic operations involved. This number increases

as the specifications of the filter are tightened.

- Digital filters are subject to round-off noise encountered in computation and if the input signal was sampled to Analogue to Digital Conversion noise.
- The design of digital filters is a far more complex task than designing an analogue filter. Computer aided design techniques in the right hands however do help to overcome this problem. Also once designed the system is usable with little or no modification for other different digital signal processing tasks. [2]

III. ARCHITECTURE DESCRIPTION

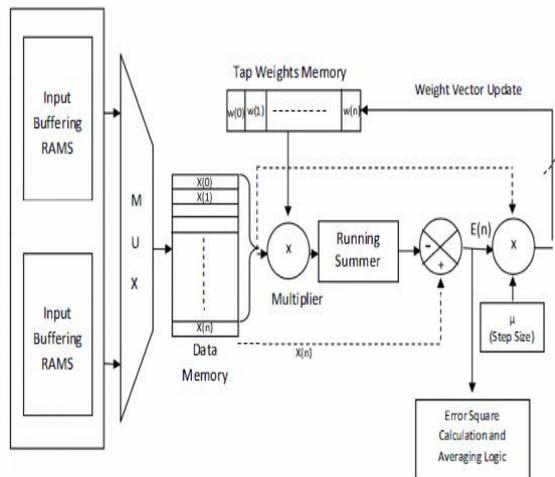


Figure 4: Block Diagram of the block LMS Architecture

A. The Weakness of the LMS Algorithm

The LMS Algorithm has one profound weakness, and it is that computational cost increases to an undesirable level as the length of the impulse response increases. This is mainly due to the fact that the algorithm lies in the time domain, leaving the algorithm at an obvious disadvantage when an impulse response is very long. The computational power required simply becomes too high for efficiency of use. There are block LMS versions of the LMS algorithm, which attempt to compensate for this problem in the time domain, however Frequency domain adaptive filtering holds the key to the solution of the very long impulse response problem.

Disadvantages:

- Convergence rate of LMS is very poor.
- BLMS algorithm provides the fastest convergence rate but highly computational expensive.
- MSE error is very high for any given SNR ratio.

The recent digital transmission systems impose the application of channel equalizers with short training time and high tracking rate. These requirements turn our attention to adaptive algorithms, which converge rapidly. The demand for fast convergence and less MSE level cannot be met by conventional adaptive filtering algorithms such as LMS. The best choice is the Block recursive least squares (BRLS) algorithm. Block Recursive Least Squares algorithms are known to exhibit better performances.

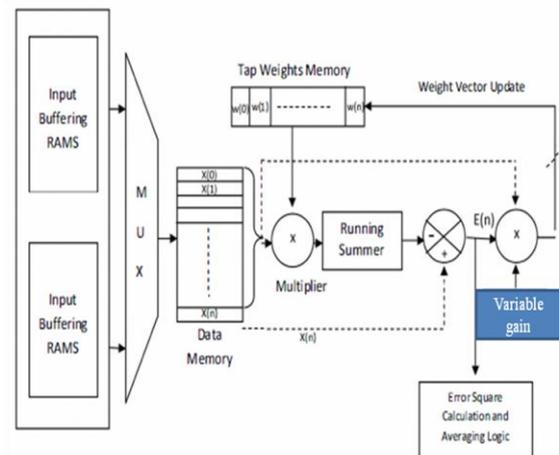


Figure 5. Block Diagram of the Proposed Architecture

The block diagram implementation is as shown in the figure5 The proposed architecture is designed to perform standalone implementation and by adding appropriate buffering blocks the architecture is made capable for real time implementation. The incoming continuous data is stored in input buffering RAMS to provide for the calculations involved until the weight updation. The data is read out alternatively from each of the input RAM blocks and is written into input data memory. Initial tap weights are Stored in tap weight memory block. Both the data and weights are readout of the memory simultaneously and passed to the multiplier which multiplies the data vector and weight vector and passes the data on to the running summer block which adds up the multiplier output .And then the running summer output, equivalent to the FIR filter output, (which ideally requires N multipliers depending on the number of taps which here is reduced to one) is then subtracted from next input sample to estimate the error. The error value obtained is then multiplied with variable gain factor and data vector to calculate the weight updates required reducing the error from each calculation. The weight updates are then added to the previous weight values and written back into the weight vector memory. The updated weights are now ready for next set of data. As mentioned earlier ideally FIR filter structure requires N multiplier depending on the number of weights considered for implementation, but with the architecture proposed. the number of multipliers is reduced to one with the pipelining of the data. The architecture presented consumes minimal hardware making it suitable for FPGA implementation.

Advantages:

- Extremely fast convergence.
- RLS algorithm to achieve the steady state error in very less no of iteration along with mean square error is also less compare to RLS algorithm for different signal to noise ratio.
- Power consumption is reduced along with hardware reduction by reducing number of iteration.

B. Performance Measures in Adaptive Systems:

1. Convergence Rate

The convergence rate determines the rate at which the filter converges to its resultant state. Usually a faster

convergence rate is a desired characteristic of an adaptive system. Convergence rate is not, however, independent of all of the other performance characteristics. There will be a tradeoff, in other performance criteria, for an improved convergence rate and there will be a decreased convergence performance for an increase in other performance. For example, if the convergence rate is increased, the stability characteristics will decrease, making the system more likely to diverge instead of converge to the proper solution. Likewise, a decrease in convergence rate can cause the system to become more stable. This shows that the convergence rate can only be considered in relation to the other performance metrics, not by itself with no regards to the rest of the system.

2. Minimum Mean Square Error

The minimum mean square error (MSE) is a metric indicating how well a system can adapt to a given solution. A small minimum MSE is an indication that the adaptive system has accurately modeled, predicted, adapted and/or converged to a solution for the system. A very large MSE usually indicates that the adaptive filter cannot accurately model the given system or the initial state of the adaptive filter is an inadequate starting point to cause the adaptive filter to converge. There are a number of factors which will help to determine the minimum MSE including, but not limited to quantization noise, order of the adaptive system, measurement noise, and error of the gradient due to the finite step size.

3. Computational Complexity

Computational complexity is particularly important in real time adaptive filter applications. When a real time system is being implemented, there are hardware limitations that may affect the performance of the system. A highly complex algorithm will require much greater hardware resources than a simplistic algorithm.

4. Stability

Stability is probably the most important performance measure for the adaptive system. By the nature of the adaptive system, there are very few completely asymptotically stable systems that can be realized. In most cases the systems that are implemented are marginally stable, with the stability determined by the initial conditions, transfer function of the system and the step size of the input.

5. Filter Length

The filter length of the adaptive system is inherently tied to many of the other performance measures. The length of the filter specifies how accurately a given system can be modeled by the adaptive filter. In addition, the filter length affects the convergence rate, by increasing or decreasing computation time, it can affect the stability of the system

IV. SIMULATION REPORTS

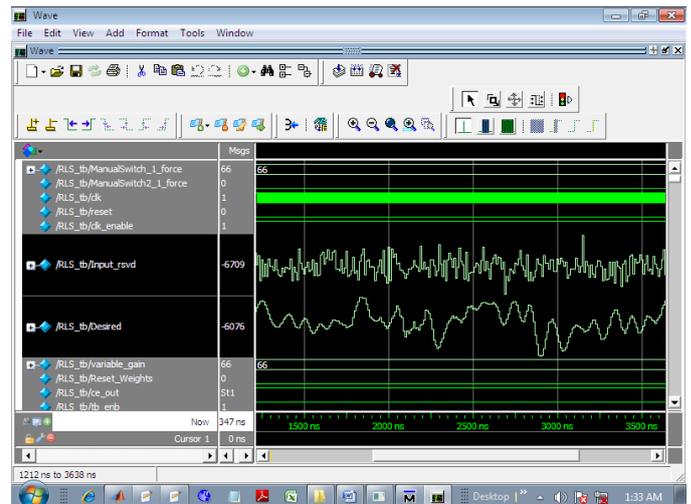


Figure 6. Simulation results

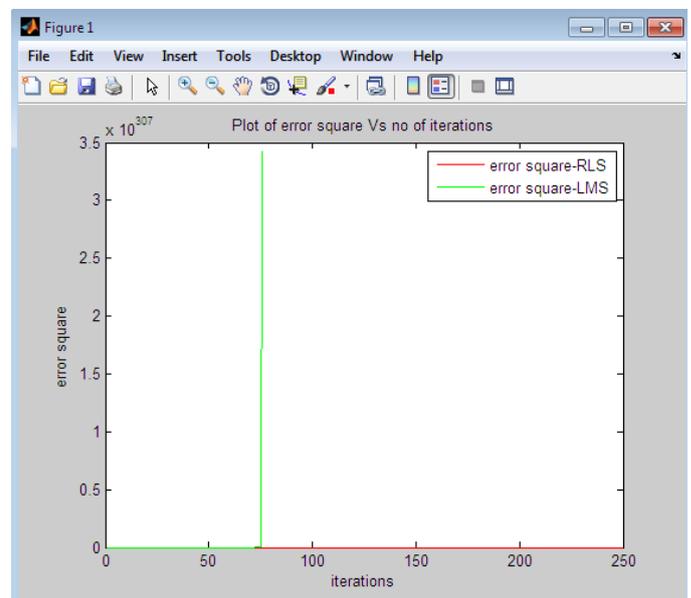


Figure 7. Variation gain Vs step size

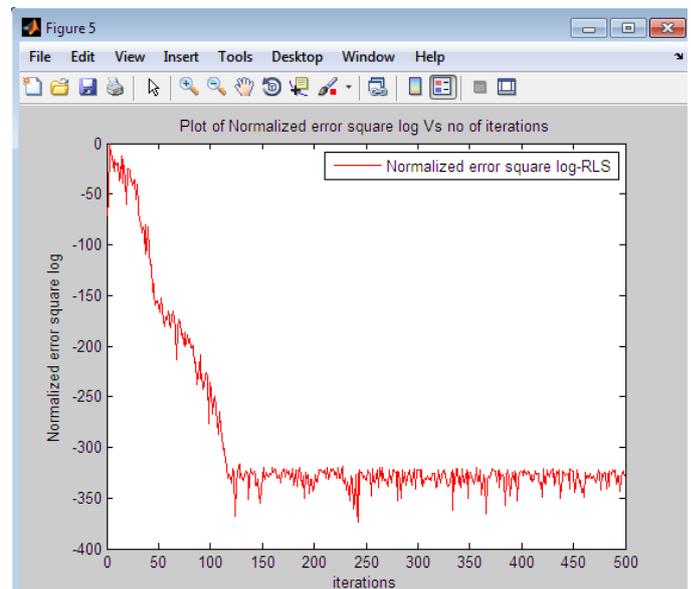


Figure 8. Learning Curves for RLS

V. DEVICE UTILISATION SUMMERY

Flow Summary	
Flow Status	Successful - Thu Aug 23 18:16:52 2012
Quartus II Version	11.0 Build 208 07/03/2011 SP 1 SJ Web Edition
Revision Name	qa
Top-level Entity Name	RLS
Family	Cyclone III
Total logic elements	5,343 / 10,320 (52 %)
Total combinational functions	5,330 / 10,320 (52 %)
Dedicated logic registers	528 / 10,320 (5 %)
Total registers	528
Total pins	69 / 183 (38 %)
Total virtual pins	0
Total memory bits	0 / 423,936 (0 %)
Embedded Multiplier 9-bit elements	46 / 46 (100 %)
Total PLLs	0 / 2 (0 %)
Device	EP3C10F256C6
Timing Models	Final

Figure9. Device utilization report

VI. CONCLUSION

The BRLS algorithm is introduced as the main adaptive algorithm in the time domain and its operation is examined. An alternative representation of signals in the frequency domain is then introduced, which allows the convolution of two signals to be calculated in a much more efficient manner. The cost of transforming the signals to and from the frequency domain must be accounted for however and for short filter impulse responses it is too high to allow frequency domain filtering replace time domain filtering. Substantial savings can be made however as the impulse response increases, an approximate crossover point was portrayed. In this report a possible implementation of the Block RLS in the Verilog HDL programming language is presented and a possible application as an adaptive equalizer explored.

REFERENCES

- [1] "Memory Based Architecture To Implement Simplified Block LMS Algorithm On FPGA"IEEE Transaction by Jayashri R, Chitra H, Kusuma S
- [2] "Communication Systems" 4th edition, Simon Haykin
- [3] IEEE Communications Magazine, March 1982, Adaptive Equalization
- [4] "Digital Signal Processing: A practical Approach" 2nd edition, E.Ifeachor and B.Jervis, Prentice Hall.
- [5] T.Yoshida, Y.Liguni, H.maeda "Systolic array implementation of block LMS Algorithm", Proc of IEEE Electronic letters 1998.
- [6] IEEE Transactions on Signal Processing, Vol. 39, No. 10, October 1991. On the Complexity of Frequency Domain Adaptive Filtering, Neil K. Jablon
- [7] "Adaptive Filter Theory", Simon Haykin
- [8] <https://www.ee.calpoly.edu/~jbreiten/C/>
- [9] <http://www.ee.nuigalway.ie/tech/misc.html>, Department of Electronic Engineering, NUI, Galway

New Cost-Sensitive Model for Intrusion Response Systems Minimizing False Positive

Jalal Baayer¹, Boubker Regragui²

^{1,2} SIME Laboratory, ENSIAS, Mohammed V Suissi University, B.P. 713 Rabat, Morocco

ABSTRACT :Dynamic Cost-sensitive Intrusion Response System (Dy-COIRS) is considered as one of the challenging intrusion response model in intrusion detection and prevention systems (IDS/IPS) field. This type of intrusion response system is faced to the issue of false positive responses (FPR) such as an error responses toward normal activities that does not affect the integrity, confidentiality availability and authentication of computer systems. This leads to high overhead which harms severally the overall network performances. In this paper, we propose an intelligent automatic dynamic cost sensitive model intended for IRS related to IDS/IPS field where the impact of false positive responses are minimized. This FPR reducing is based on an algorithmic approach with a linear model theory.

Keywords: Dynamic Cost-sensitive Intrusion Response System (Dy-COIRS), IDS (Intrusion Detection System), IPS (Intrusion Prevention System), false positive responses, linear model.

I. INTRODUCTION

Dy-COIRS is able to assure responses against attacks which are generally divided into the following four categories [1], [13]: Denial of service, User to root (U2R), Remote to local (R2L) and Probe. The decision of response launch is based on cost approach without the aid of manual interventions of a network administrator. A Dy-COIRS is seen as an automatic intelligent system because its components can dynamically establish a release in the form of response further to intrusions analysis.

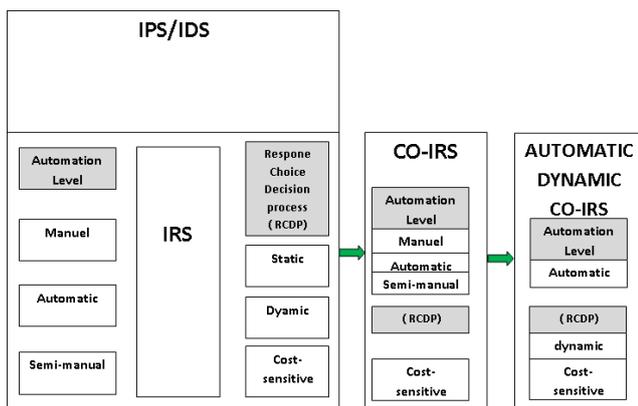


Fig. 1. Dy-COIRS used with IPS / IDS Systems

In a Dy-COIRS, the response can be given dynamically for IDS or IPS systems, by considering two values: intrusion affectation damage and response deployment cost.

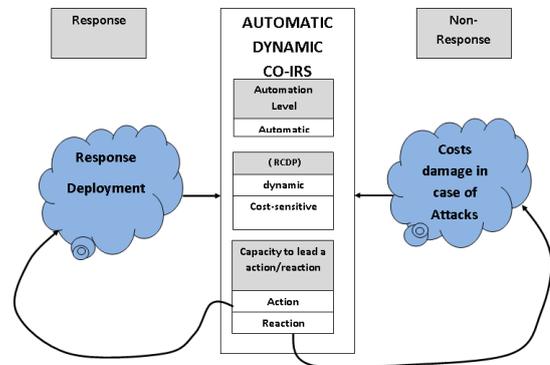


Fig. 2. The Functioning of Dy-COIRS

Those costs values are analysed to decide on the necessity of an optimal given response [2]. So, the success of a given response is strongly dependent on the good balance between the attack affectation damage and the system resources restoring costs.

In practice, the IRS has a false positive response when it lanches a wrong response against a real attack or in front of a normal activity. False positive responses in Dy-COIRS can seriously affect, disrupt the efficiency and harshly degrade the overall performances of IDS/IPS [3] [4] [12].

When the most known intrusions can be avoided or countermeasured by appropriate responses, false positive responses are still subject of various research works. This fact is due to their presence in IRS and Dy-COIRS field as “mistake” phenomena related to an abnormal action in front of an innocent behaviour. So, to decrease the impact of these false positive responses in IDS/IPS field, many reducing models concepts were developed [6] [7] [8] [9] [10]. Indeed, these false positive reducing models permit to report the real intrusions and attacks with false positive responses minimization. Consequently, given appropriate responses against real intrusions increase the operation quality and the accuracy of such IRS and specially Dy-COIRS.

In a Dy-COIRS, the fixed cost model is pre-discussed and related to the cost of each intrusion and the cost of the response launched as countermeasure [11]. In reality, the choice of a model with a big enough cost of false positive responses generates a large extra cost which clearly undermines the performance and efficiency of an Dy-COIRS. So, to avoid this problem a new cost-sensitive model for intrusion response systems is required.

In this paper, we present a new intelligent cost-sensitive model for intrusion response systems used for IDS/IPS to limit and minimize the impact of false positive responses. Our proposed model is based on an algorithmic approach with a linear model theory and on cost approach of intrusions and responses. Moreover, our proposition of false positive minimization enables to network administrators to limit and reduce the cost generated by a

false positive response and increase the performance of Dy-COIRS.

The rest of this paper is organized as follows. Section 2 presents a related work that gives an overview on Dy-COIRS and false positives responses minimization related to intrusions responses systems. Section 3 presents models with mathematical approach and especially linear models. Section 4 presents our improvement. Section 5 presents simulations and results. The conclusion is given in the last section 6.

II. RELATED WORK

The domain of Dy-COIRS with FPR minimization was not enough targeted by the researchers these last years as it was generally with intrusions detection, prevention or response systems. But, there are few numerous research works in this domain which can be revealing.

In the field of Dy-COIRS, many research works are done independently of the purpose of FPR minimization. We can mention the following researchers works: B. Foo [14], T. Toth [2], I. Balepin [15], M. Jahnke [16], S. Yu [17], M. Papadak [18], K. Haslum [13], C. P. Mu [19], W. Kanoun [20] and N. Kheir [21]. All these works present different models of Dy-COIRS without the explicit false positive responses minimization.

At the first, Denning asserted in his study [22], that the study of the costs is not seen in the aspect of an authentic knowledge. This work switched on the first light around the notion of cost in the field of intrusions responses. Northcutt came to treat in [23] the methodology of studies of the risks in the computer systems by describing measures basing itself on degrees of criticality and destruction. Approach proposed by Balepin [15] basing on the principle of the representatives of services used a graphic prototype for the selection of the optimal responses with the institution of a typical hierarchy of resources by engendering the maximum of privileges with the minimum of cost. Toth [2] used of a prototype of a computer network by taking into account means (functions/ services), the users, the type of the network and the control access systems. The costs of the responses are measured by basing on the decrease of the values of the capacities of the resources. The work of followers [14] was presented as Framework allowing the choice and the deployment of the automatic response against intrusions basing on two categories of graphic plans: a plan of service and plan of response. These models and these solutions evoked above are not coherent between them. Every proposal has a concept of evaluation and selection of response cost sensitive with a different vision. The works of [22] and [23] evoked for the first time the notion of cost, and the method of study of the risks in the field of detection of intervention without treating the response cost with connection with intrusion cost. The works presented by [2] and [20] considered the response cost in contribution with the resources of system, by showing several processes of estimation. Toth [2] calculates the cost of response as a function of decrease of capacity of system. [15] measured the cost of answer being the sum of the costs manually committed by the affected system resources. All works done by [2], [13], [14], [15], [16], [17], [18], [19], [20],[21], [22] and [23] did not take into account the impact of cost false positive response, and his necessary minimization to have appropriate responses.

Other side, in the field of false positive responses minimization, many research works are done independently of the purpose of Dy-COIRS field. We can mention the following researchers works: Subramanian [6], Benjamin [7], Emmanuel Hooper [8], Hassen Sallay [9] and Kai Hwang [10]. Subramanian [6] introduced a Content Split Approach (CSA), made particularly for the signature databases related to network intrusion detection and prevention to reduce false positive responses. Benjamin [7] suggested a relationship of the information associated to the characteristics of the supervised information system, information about the faults, information security utilities of information security used to supervision events. Hooper [8] proposed a model to minimize false positives responses based on adaptive responses of firewall rules. This model represents a mixture of firewall structural design connected with response rules, to reject entrance to crucial segments to doubtful hosts in the computer network. Hassen Sallay [9] reflected on a scalable structure for IDS shared for networks with a large flow to amend efficiency. Kai Hwang [10] suggested a hybrid model of signature based Intrusion Detection System IDS and Anomaly Detection System (ADS), to reduce false-positive and to detect unknown intrusions.

There are little research works that targeted the minimization of false positive responses within the Dy-COIRS. We can mention the following researchers works: W. Lee[1], Strasburg [24], S. Tanachaiwiwat [25], and N. Stakhanova [5]. Work by Lee [1] considered experimental costs of the effects of the intrusions and the measures taken by the responses as criteria for the responses choice against the intrusions already classified. This work introduces a cost-benefit measure which incorporates multiple dimensions of cost in the face of an intrusion: response cost. This work is done only for IDS and not IPS. It did not target the Dy-COIRS. And the minimization of the cost of False Positive was treated usually implicitly with Consequential Cost Reducing and not explicitly as specific False Positive minimization. The works of Strasburg [24] and N. Stakhanova [5] presents a host-based framework for cost-sensitive intrusion response selection with a method for evaluating each intrusion response with respect to the risk of potential intrusion damage, effectiveness of response action and response cost for a system. The minimization of false positive response is treated implicitly with damage cost optimization. S. Tanachaiwiwat [25] presented a framework constructed with three essential modules: the IDS (intrusion detection system), the RAS (risk assessment systems), and the IRS (intrusion response system). The RAS is able to distinguish different kinds of false alarms or miss detections. The minimization of false positive responses is made in the context of the risk assessment systems without a using of specific algorithm for FP reducing.

III. LINEAR CONTROL SYSTEM

In the mathematics sense, a system is linear if we can apply it the principle of superimposing. Basing on a physical point, linear system can be defined more restrictive as a system that is can be described by differential equations with finished order and constant coefficients. With this definition, we can associate with the system, by means of the transformed of Laplace, a transmittance $H(p)$

which is a rational fraction with $p=j\omega$. In automatic, we complete frequently the of linearity with the transmittance associated with the pure delay, that is a term of the form $\exp(-\alpha p)$ with α is a constant time. The methods of study of the linear systems are very powerful in reason of the available tools (linear algebra, differential equations and differential systems).

The linear systems are relatively simple from a mathematical and algorithmic point of view, and purely exactly because of the linearity of the equations.

An linear control system is a commanded system possessing a device of return allowing to compensate for the infidelity of a physical system. It includes:

- **The direct chain H(p):** it is the commanded system which is subjected to the influence of the disturbances and thus miss of fidelity. Its transmittance is often noted H (p).
- **The chain of return K:** it converts the greatness of exit in a tension which is the signal of return x_r . This sensor must be faithful, so, insensible with the disturbances.
- **The organ of display K:** it transforms the wished value Y_e of y (instruction) to tension x. it is not present in all control process.
- **The comparator:** it elaborates the signal of error $e = x - x_r$

A linear control system in a certain range around the point of rest, provided with an organ of display, a buckle of return and supposed initially in the rest, has the following functional plan:

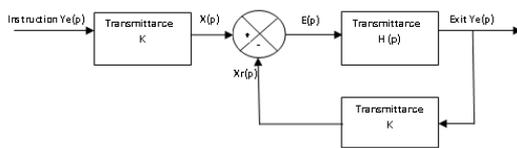


Fig. 3. Functional plan of linear control system

IV. OUR IMPROVEMENT

Generally, each activity in the computer network has two states: real attack or normal activity. Each such Dy-COIRS can have three reactions: no response, true response or false response. We can present all possible combinations between intrusions cases and responses cases on the following figure:

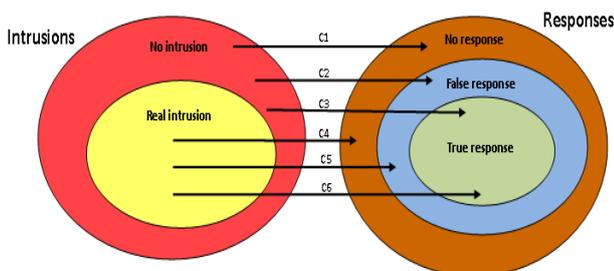


Fig. 4. The Reaction cases of Dy-COIRS

The false positive response is a simple wrong countermeasure which can be launched by IDS or IPS through IRS and specially Dy-COIRS when a legitimate activity is considered wrongly as attack.

The Figure 4 illustrates that the false positive response corresponds to the two cases: C2 and C5. False negative response corresponds to the case C4. True response corresponds to the two cases: C1 and C6. The case C3 is not feasible because it cannot logically happen in reality.

This problem of false positive responses can be treated at two basic approaches. The first approach is the conception of the IDS and IPS. The second is the implementation part of these systems. Indeed, for a given legitimate activity in the network, the IDS or IPS intercepts it as an attack, and sent an order to the IRS to launch immediately a response independently to any estimation to the cost or the impact of this countermeasure on the actual environment. So the intrusion detecting processing shown with conception and implementation approach of IDS or IPS is not enough to limit the impact of this phenomenon of false positive on network performance and efficiency security strategy.

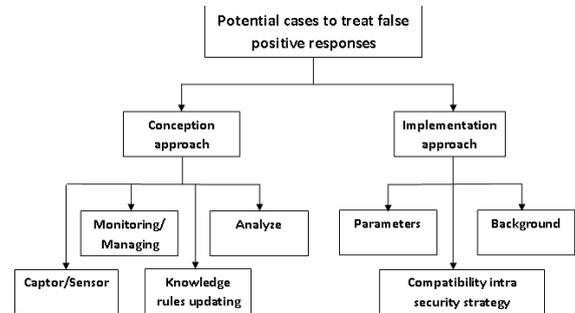


Fig.5. Typical cases of false positive treating

The Figure 5 illustrates typical cases and approaches of treating of false positive responses. In the first step, for a given innocent activity in the network, IDS or IPS considers it wrongly as intrusion at design or operational approach. In second step and after making mistake, IDS or IPS send an order to Dy-COIRS to launch an passive response as alarms or an active response as blocking of the traffic considered as doubtful or stopping packages every time an attack is discovered. As a result, all passive or active false positive responses that verify the classic conditions detection through conception and operational approaches present an overhead which impact directly the network performance and the IDS or IPS severity.

Recent works [1], [5], [24] and [25] are still using the minimization of false positive responses implicitly with damage cost optimization and in some cases it is made in the context of the risk assessment systems without a using of specific algorithm for FP reducing. This choice of classic process give a greatest weakness related to the cost minimizing of false positive. Indeed, without this cost model for false positive, administrators will have difficulty to understand the cost of false positive in their organizations, depending on the way they implement their IDS, IPS and Dy-COIRS. Thereafter, without this model, we have not an effective tool to help network administrators to describe quantifiable costs of false positives responses and other costs that are concrete, still difficult to quantify such as loss of services.

In this section, we present a cost sensitive responses model aiming to limit the impact of false

positives responses launched wrongly by Dy-COIRS in the case of an innocent activity in the network. Our approach has the advantage not to require any complex process because it only based on the linear system theory. Our cost model presents the responses costs launched against attacks and intrusions threatening a network during a period and this model helps for minimizing of false positive responses impact. Moreover, this responses costs presentation and false positive responses impact minimizing is a lightweight calculation and it is based on the standard theory of linear systems. The choice of the linear approach is due to its simplicity

Referring to the Figure 5, the phenomena of false positive appears when a IDS or IPS detect wrongly a legitimate activity as an attack. This mistake aspect can appear at design or implementation approach. So, an error response as false positive should be launched by Dy-COIRS. That means that minimizing false positives responses can be done by reducing their responses costs. Based on this observation, we can define the cost response results of an Dy-COIRS relates to IDS or IPS through the Cost Response Matrix

$$M = (c_{ij}) \quad (1)$$

M is a square matrix of order n+1, where n is the number of different intrusion types. For $1 < i, j < n$, the i-th row corresponds to intrusion type i. The j-th column corresponds to response type j launched by The Dy-COIRS. The matrix element (m_{ij}) is equal to the cost represented by response type j against the attack type i. The intrusions are indicated by the rows and responses by columns.

The costs false negative responses correspond to all miss detection at the last column marked as the FNR domain. The cost false positive response are at bottom row with $i = n+1$ for no attacks marked as the FPR domain.

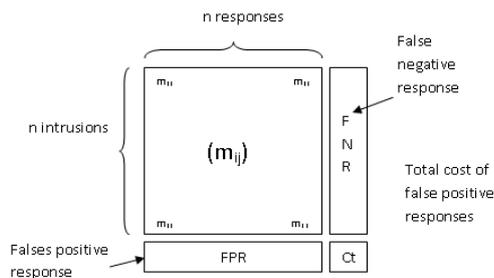


Fig.6. M cost Response matrix

In the following part of this work, in order to show the importance of our proposed, we define the matrix A as the initial cost matrix of cost related to the Dy-COIRS. It represents the knowledge module of cost rules using to evaluate the cost response. Each (a_{ij}) show the cost deployment of the response j against the intrusion i. The values of (a_{ij}) are evaluated by technical and financial administrators using a real journal of history of events and specially different intrusions. The costs of different responses deployed are calculated and noted in this journal. The matrix A is a represents cost rules module defined for each Dy-COIRS.

Indeed, We consider also, the matrix B representing the number of responses launched against intrusions following the the random low. This matrix will be constructed after a long observation of system targeted by intrusions during the study period. Each (b_{ij}) illustrates the number of the response of type j against the intrusion of type i.

For each intrusion of type i, the total number of responses launched against it is represented by Ni:

$$N_i = \sum_{k=1}^{k=n} b_{ik} \quad (2)$$

Where:

- b_{ik} is the number of the response of type k against the intrusion of type i.
- n is the total number of intrusions and responses

The probability that a response of type k, will be launched against the intrusion of type i is represented as following:

$$\pi_{ik} = (b_{ik} / \sum_{m=1}^{m=n} b_{im}) \quad (3)$$

We define also the matrix C=(c_{ij})= (π_{ij}) that represents the probability matrix of responses launched against intrusions . The probability matrix is described as following in the proposed model:

Fig.7. C probability matrix of responses

Thereafter, define D=(d_{ij}) as the matrix of real response cost or the normalized cost matrix .It is observed during period study. We can define the matrix D as a multiplication of the two matrixes A and C.

$$D = A \times C \quad (4)$$

Each d_{ij} is obtained as the cost of response j launched against an intrusion i, following the term:

$$d_{ij} = \sum_{k=1}^{k=n} a_{ik} \times c_{kj} \quad (5)$$

For our proposed algorithm for the minimization of false positive, we consider the definitions bellow:

- *Tolerance value (TV)*: a value chosen for each IDS/IPS, under it, the rate of false positive allows to have an optimal work of IDS/IPS.
- *Cfp*: false positive responses cost defined following the matrix D.
- *Ctotal*: responses total cost defined following the matrix D.
- *Rt*: false-Positive Cost Ratio= Cfp / Ctotal
- *Rav*: the average false-Positive Cost Ratio
- Matrix A_0 : Present the initial values of the matrix A_0 related to each IDS/IPS/ Dy-COIRS.

Our proposed algorithm for the minimization of false positive is presented as following:

Matrix A = Matrix A_0 ;

TV=TV₀;

Step 1: Matrix B → Matrix C;

D=A×C;

Cfp= $\sum_{i=1}^{i=n+1} d_{i, n+1}$;

Ctotal= $\sum_{i=1}^{i=n+1} \sum_{j=1}^{j=n+1} d_{ij}$;

If (Rt < TV) then (our IDS/IPS is working optimally : OK);

If not (Rt ≥ TV) then (our IDS/IPS is working non optimally: NOK):

```

{ Rav= Rt/(n+1);
  For i=1 to i=n+1;
    If (aij ≥ Rav) then
      aij=aij - 1 ;
    EndIf;
  EndFor; }
    
```

EndIfnot;

Go to step1;

V. SIMULATION AND RESULT

For illustrative purpose, we tested our model on real IDS with simulation experiments over the following alarm matrix resulted from data entries extracted from an IDS evaluation report by IT maintenance dept. in Q1/2012. This matrix corresponds five 5 attacks types in the following Table. The matrix entries correspond to a 3 month-long monitoring of a troubleshooting and maintenance entity.

Cost matrix						Matrix Random Generation						Gi					
		16	0	2	0	0	22			4	4	7	10	3	4	32	
		0	13	0	1	0	10			2	4	5	7	1	5	24	
	A	2	0	41	3	0	19		B	6	3	9	4	4	2	28	
		0	0	0	8	0	24			1	7	8	3	3	3	25	
		0	0	0	0	2	4			8	4	1	5	3	2	23	
		5	1	4	5	0	0			3	3	4	5	3	6	24	
Cost matrix (multinomiale law)						The probability model:											
		5,2	4,96	7,81	9,9	4,5	7,6			0,13	0,13	0,22	0,31	0,09	0,13	1	
		2,4	3,7	4,7	6	1,9	5,3			0,08	0,17	0,21	0,29	0,04	0,21	1	
	Ax C=	D	12	7,86	17,7	11	8,8	8,3		C	0,21	0,11	0,32	0,14	0,14	0,07	1
			3,3	5,24	6,56	6	4	7			0,04	0,28	0,32	0,12	0,12	0,12	1
			1,2	0,85	0,75	1,3	0,8	1,2			0,35	0,17	0,04	0,22	0,13	0,09	1
			1,8	2,62	4,19	3	1,7	1,7			0,13	0,13	0,17	0,21	0,13	0,25	1

Fig.8. Our first simulation of the cost model of responses

We define the knowledge module that can compare the False-Positive cost Ratio with a tolerance value (TV).

- Step 1 : If the 8% is less than the TV our IDS/IPS is working optimally → OK
- Step 2 : If the 8% is more than the TV our IDS/IPS is working non optimally → NOK
 - We have to react to this situation by minimizing the false-Positive Cost Ratio cost line
 - The average of the False-Positive cost is = 15/6 = 2.5
 - If the Cost is > 2.5 → put Cost - 1
- Goto Step 1

Cost matrix						Matrix Random Generation						Gi					
		16	0	2	0	0	22			7	5	4	6	4	9	35	
		0	13	0	1	0	10			1	6	9	8	1	3	28	
	A	2	0	41	3	0	19		B	4	3	4	3	4	3	21	
		0	0	0	8	0	24			8	1	10	7	2	1	29	
		0	0	0	0	2	4			6	1	8	1	10	4	30	
		4	1	3	4	0	0			1	6	2	7	10	4	30	
Cost matrix (multinomiale law)						The probability model:											
		4,3	6,97	3,68	8,2	9,5	7,3			0,20	0,14	0,11	0,17	0,11	0,26	1	
		1,1	4,82	5,19	6,3	3,9	2,8			0,04	0,21	0,32	0,29	0,04	0,11	1	
	Ax C=	D	9,7	10	10,3	11	15	9		C	0,19	0,14	0,19	0,14	0,19	0,14	1
			3	5,08	4,36	7,5	8,6	3,5			0,28	0,03	0,34	0,24	0,07	0,03	1
			0,5	0,87	0,8	1	2	0,8			0,20	0,03	0,27	0,03	0,33	0,13	1
			2,5	1,35	2,73	2,4	1,3	1,7			0,03	0,20	0,07	0,23	0,33	0,13	1

Fig.9. Our second simulation minimizing cost false positive

If we suppose that the TV = 7.5% and as we decrease the False-Positive Cost value by 1, we have u = 12 false-positive alarms at the bottom row. Our total cost is 182 that mean our false-Positive Cost Ratio is 7% which means that our IDS/IPS is now with an optimal Dy-COIRS.

VI. CONCLUSION AND PERSPECTIVES

In this paper, we propose a cost-sensitive model for Intrusion Response Systems to limit the impact of false positive in IRS and specially Dy-COIRS. Our approach reduces the cost of false positive responses launched by Dy-COIRS in the case of an innocent activity comparing to the classic methods based on detection without cost approach revealing defined at the beginning of a communication. Therefore, the cost of false positive responses are reduced which increase the IRS and COSIRS performance. As a future work, we plan to study the COSIRS performance in the case where we combine our cost model, based on linear system theory, with another model minimizing the false negative responses.

REFERENCES

- [1] W. Lee, W. Fan, M. Millerand, S. Stolfo, and E. Zadok. Toward cost-sensitive modeling for intrusion detection and response. *In Journal of Computer Security, volume 10, pages 5–22, 2002.*
- [2] T. Toth and C. Kregel, Evaluating the impact of automated intrusion response mechanisms, *In*

- proceeding of the 18th Annl Computer Security Applications Conference, Los Alamitos, USA, 2002.*
- [3] K. Timm, *Strategies to reduce false positives and false negatives in NIDS*, Security Focus Article, available online at: <http://www.securityfocus.com/infocus/1463>, 2009
- [4] M.J. Ranum, *False Positives: A User's Guide to Making Sense of IDS Alerts* (ICSA Labs IDSC, 2003).
- [5] N. Stakhanova, S. Basu, and J. Wong. A cost-sensitive model for preemptive intrusion response systems *In Proceedings of the IEEE AINA, pages 428–435*, 2007.
- [6] Subramanian Neelakantan et. al. (2009) "Content-Split Based Effective String-Matching for Multi-Core Based Intrusion Detection Systems", First International Conference on Computational Intelligence, Communication Systems and Networks Pages: 296-301 ISBN:978-0-7695-3743-6
- [7] Benjamin Morin 1, Ludovic M, Herv Debar, and Mireille Ducass (2007) *M2D2: A Formal Data Model for IDS Alert Correlation, Volume 2516/2002, pages 115-137*
online <http://www.springerlink.com/content/cwp428tlhf35rwnba/>
- [8] Emmanuel Hooper (2006), "An Intelligent Intrusion Detection and Response System Using Network Quarantine Channels: Adaptive Policies and Alert Filters", *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2006 Workshops)(WI-IATW'06)*, pp. 16-21, 0-7695-2749-3/06 \$20.00 © 2006.
- [9] Hassen Sallay, Khalid A. AlShalfan, Ouissem Ben Fred, (2009), "A scalable distributed IDS Architecture for High speed Networks", *IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.8, August 2009*.
- [10] Kai Hwang, MinCai, Ying Chen and Min Qin "Hybrid Intrusion Detection with Weighted Signature Generation over Anomalous Internet Episodes"(2007), *IEEE Transactions On Dependable And Secure Computing, Vol.4, No.1, January-March 2007*, accessed on 22.02.08, <http://ieeexplore.ieee.org/search/wrapper.jsp?arnumber=4099191>.
- [11] J. Baayer, B.Regragui- *WOTIC'09* – " Architecture Fonctionnelle d'un IPS, Etat de l'Art et Classification de Ses Systèmes de Réponse d'Intrusion (IRS)", *December 25 2009*, Université Ibn Zohr, Agadir, Morocco
- [12] Stefano Zanero(2007), *Flaws and Frauds in the Evaluation of IDS.IPS Technologie, first accessed on 21.09.07*, <http://www.first.org/conference/2007/papers/zanero-stefano-paper.pdf>
- [13] K. Haslum, A. Abraham and S. Knapskog, "DIPS: A framework for distributed intrusion prediction and prevention using hidden markov models and online fuzzy risk assessment," *In 3rd International Symposium on Information Assurance and Security*, pp. 183-188, Manchester, United Kingdom, 2007.
- [14] B. Foo, Y.-S.Wu, Y.-C. Mao, S. Bagchi, and E. H. Spafford. ADEPTS: Adaptive intrusion response using attack graphs in an e-commerce environment. *In Proceedings of DSN*, pages 508–517, 2005.
- [15] I. Balepin, S. Maltsev, J. Rowe, and K. Levitt. Using specification-based intrusion detection for automated response. *In Proceedings of RAID*, pages 136–154. Springer, 2003.
- [16] M. Jahnke, C. Thul, and P. Martini. Graph based metrics for intrusion response measures in computer networks. *In Proceedings of the IEEE LCN*, pages 1035–1042, 2007.
- [17] S. Yu and Z. Rubo, "Automatic intrusion response system based on aggregation and cost," *in International Conference on Information and Automation (ICIA)*, 2008, pp. 1783-1786.
- [18] M. Papadaki and S. M. Furnell, "Achieving automated intrusion response: a prototype implementation," *Information Management and Computer Security*, vol. 14, no. 3, 2006, pp. 235-251.
- [19] C. P. Mu and Y. Li, "An intrusion response decision making model based on hierarchical task network planning," *Expert systems with applications*, vol. 37, no. 3, 2010, pp. 2465-2472.
- [20] W. Kanoun, N. Cuppens-Boulahia, F. Cuppens and S. Dubus, "Risk-Aware Framework for Activating and Deactivating Policy-Based Response," *Fourth International Conference on Network and System Security*, pp. 207-215, 2010.
- [21] N. Kheir, N. Cuppens-Boulahia, F. Cuppens and H. Debar, "A service dependency model for cost sensitive intrusion response," *Proceedings of the 15th European Conference on Research in Computer Security*, pp. 626- 642, 2010.
- [22] D. Denning. *Information Warfare and Security* (Addison-Wesley, 1999).
- [23] S. Northcutt. *Intrusion Detection: An Analyst's Handbook*. (New Riders Publishing, 1999).
- [24] C. Strasburg, N. Stakhanova, S. Basu and J. S. Wong, "A Framework for Cost Sensitive Assessment of Intrusion Response Selection," *Proceedings of IEEE Computer Software and Applications Conference*, 2009.
- [25] S. Tanachaiwiwat, K. Hwang and Y. Chen, *Adaptive Intrusion Response to Minimize Risk over Multiple Network Attacks* (ACM Trans on Information and System Security, 2002).

Review of WI-FI Security techniques

Promila¹, Dr. R. S. Chhillar²

*(Department of Computer Science and Application, M. D. U. Rohtak, India)

** (Department of Computer Science and Application, M. D. U. Rohtak, India)

Abstract: *Wireless technology provides us many benefits like portability and flexibility, increased productivity, and lower installation costs. Wi-Fi networks can be accessed with laptops, mobile phones, cameras, game consoles, and an increasing number of other consumer electronic devices. Wireless technologies have become increasingly popular everyday in business as well as in personal lives. Wireless Networking changed completely the way people communicate and share information by eliminating the boundaries of distance and location. In this paper we are discussing about the wireless network challenges and IEEE 802.11 Standards and WEP protocol.*

Keywords: WI-FI, WEP, SSID, MAC, WiMAX, DoS .

I. Introduction

Wi-Fi is the name of the popular wireless networking technology that uses radio waves to provide wireless high-speed internet and network connection. The Wi-Fi alliance, the organization that owns the wi-fi (registered trade mark) term specifically defines Wi-Fi as any "wireless local area network (WLAN) products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." A common misconception is that the term Wi-Fi is short for "wireless fidelity," however this is not the case. Wi-Fi is simply a trademarked term meaning IEEE 802.11x. Initially, Wi-Fi was used in place of only the 2.4GHz 802.11b standard, however the Wi-Fi Alliance has expanded the generic use of the Wi-Fi term to include any type of network or WLAN product based on any of the 802.11 standards, including 802.11b, 802.11a, dual-band, and so on, in an attempt to stop confusion about wireless LAN interoperability. Wi-Fi works with no physical wired connection between sender and receiver by using radio frequency (RF) technology, a frequency within the electromagnetic spectrum associated with radio wave propagation. When an RF current is supplied to an antenna, an electromagnetic field is created that then is able to propagate through space. The cornerstone of any wireless network is an access point (AP). The primary job of an access point is to broadcast a wireless signal that computers can detect and "tune" into. In order to connect to an access point and join a wireless network, computers and devices must be equipped with wireless network adapters. Wi-Fi is supported by many applications and devices including video game consoles, home networks, PDAs, mobile phones, major operating systems, and other types of consumer electronics. Any products that are tested and approved as "Wi-Fi Certified" (a registered trademark) by the Wi-Fi Alliance are certified as interoperable with each other, even if they are from different manufacturers. For example, a user with a Wi-Fi Certified product can use any brand of access point with any other brand of client hardware that also is also "Wi-Fi Certified". Products that

pass this certification are required to carry an identifying seal on their packaging that states "Wi-Fi Certified" and indicates the radio frequency band used (2.5GHz for 802.11b, 802.11g, or 802.11n, and 5GHz for 802.11a).

II. Related work

Wireless is in everywhere like-More devices are using Wi-Fi:- Cell phones

- Digital cameras
- Printers
- PDAs
- Video game controllers
- Televisions
- Speakers
- Refrigerators etc [5].

III. Wireless Networks Challenges

Wireless Networks plays the most important role in the development of the information in between individual-to-individual, business-to-business, and individual-to-business. It changed completely the way of sharing of the information but still there are lot of challenges which are the hurdles in the wide adaptation of wireless network technology [1], [2].we have to understand the main problems that not only WI-FI network faces but all the networks faces are –CIA that is confidentiality, integrity and authentication.

Confidentiality:

Allow only the authorised person to read the encrypted messages or the information.

Integrity:

It is defined as the information not being opened by third person and it should reach in the same format as it was sent by the sending party.

Authentication:

The parties sending or receiving messages make sure that, who they say they are, and have right to undertake such actions.

The main issue in the security of wireless signal is its mode of transmission .wireless signals are transmitted through the electromagnetic waves; these waves can not be contained physically. In wireless networks the signals are communicated via air, hence can be easily intercepted with the help of right transceiver equipment.

IEEE 802.11 Standards

In 1997, IEEE ratified the 802.11 standard for WLANs. The IEEE 802.11 standard supports three transmission methods, including radio transmission within the 2.4 GHz band. In 1999, IEEE ratified two amendments to the 802.11 standard—802.11a and 802.11b—that define radio transmission methods, and WLAN equipment based on

IEEE 802.11b quickly became the dominant wireless technology [10]. IEEE 802.11b equipment transmits in the 2.4 GHz band, offering data rates of up to 11 Mbps. IEEE 802.11b was intended to provide performance, throughput, and security features comparable to wired LANs. In 2003, IEEE released the 802.11g amendment, which specifies a radio transmission method that uses the 2.4 GHz band and can support data rates of up to 54 Mbps. Additionally, IEEE 802.11g-compliant products are backward compatible with IEEE 802.11b-compliant products.[7].

IEEE Standard or Amendment	Maximum Data Rate	Typical Range	Frequency Band	Comments
802.11	2 Mbps	50-100 meters	2.4 GHz	
802.11a	54 Mbps	50-100 meters	5 GHz	Not compatible with 802.11b
802.11b	11 Mbps	50-100 meters	2.4 GHz	Equipment based on 802.11b has been the dominant WLAN technology
802.11g	54 Mbps	50-100 meters	2.4 GHz	Backward compatible with 802.11b

Summary of IEEE 802.11 WLAN Technologies [7]

WEP:-

WEP protocol is part of the IEEE 802.11 standard [3], [8], [9], [10], [11], [13]. It was introduced in 1997. WEP is used in 802.11 network to protect link level data during the wireless transmission. WEP was the first cryptographic protocol which are developed for the WI-FI to enable privacy and authentication. WEP uses the shared key authentication mechanism and is based on secret cryptographic key. WEP protocol uses the RC4 (Rivest Cipher 4) stream cipher algorithm to encrypt the wireless communications. This RC4 stream algorithm protects the contents from disclosure to eavesdroppers. WEP support 40-bit key and with extension it also support 128 or even 256 bit key also. WEP was designed to protect a wireless network from eaves dropping. WEP uses linear hash function for data integrity. In WEP there is no key management and no replay detection facility. But in 2001 several serious weaknesses were identified. Now, WEP connection can be cracked within minutes. After having such type of vulnerabilities, in 2003 the WI-FI Alliance WEP had been replaced by WPA. The main problem of WEP was-it uses static encryption keys.

WPA/WPA2:-

WPA and WPA2 are two security protocols developed by WI-FI Alliance [9], [10], [11], [13]. WPA provides developed with the purpose of solving the problems in WEP cryptographic method. WPA was developed in 2003. Both WPA and WPA2 have two modes of operation:

Personal and Enterprise. The Personal mode involves the use of a pre-shared key for authentication, while the Enterprise mode uses IEEE 802.1X and EAP for this purpose. WPA2 was introduced in September 2004. WPA addresses a subset of the IEEE 802.11i specification that addresses the weaknesses of WEP. WPA2 extends WPA to include the full set of IEEE 802.11i requirements. WPA is easier to configure and it is more secure than WEP. WPA uses the improved encryption algorithm known as TKIP (Temporal Key Integrated Protocol). TKIP provides each client with a unique key and uses much longer keys that are rotated at a configurable interval. It also includes an encrypted message integrity check field in the packets; this is designed to prevent an attacker from capturing, altering and/or resending data packets which prevent Denial-of-Service and spoofing attack. WPA can be operated with the help of RADIUS server or without RADIUS servers. Now, TKIP can be broken easily. WPA2 uses Advanced Encryption Standard. WPA2 may not work with some older network cards. WPA2 have the 4 main key factors:-

1. mutual authentication
2. strong encryption
3. interoperability
4. Ease to use.

These are the 4 main advantages of WPA2. WPA and WPA2 use the cryptographic hash function for data integrity. WPA and WPA2 both provides the key management and replay detection.

The fundamental aspect of Wireless Networks in maintaining security is to maintain Confidentiality where the receiver should receive the actual transmitted information from the sender. The message authentication provides integrity to both sender as well as receiver. The Wireless Link should be always available and should be secured from outside world like malicious attacks as well as DoS Attacks (Denial of Service Attacks).

There are basically two common attacks which compromise the security and authentication mechanism of Wireless Networks i.e. Message Reply Attack and Man in the Middle Attack. The Message reply attack acts principally on the authentication and authentication key formation protocols. The Man in the Middle Attack (MitM) attack occurs on that security mechanism which doesn't provide mutual authentication.

Various other attacks like Session Hijacking, Reflection attacks are there which affects the security mechanism of Wireless Networks.

IEEE helped in securing the wireless networks by providing the basic measures for securing wireless network and it also provide CIA factors by disabling SSID, use of MAC i.e. Media Access Control address filtering and WPA/WPS protection mechanism. The recent developments in computer technology and software developments notice that these mechanisms have network vulnerable attack. So, due to these vulnerabilities WiMax standards comes into existence, for solving the short comings of 802.11 wireless networks [4]. WiMax is the new advancement in the wireless network. WiMax is still undergoing development and still the securing problems are not being decreased by WiMax technology. It also has some drawbacks like it lack mutual authentication and is suspected to relays attacks, spoofing of MAC address of

Subscriber Station (SS) and PMK authorization vulnerabilities.

IV. Conclusion

Wi-Fi security is not an easy task. Wireless network security is more difficult than wired network security. There are many protocols or standards or we can say technologies for wireless network security but every protocol has its demerits, until now there is no protocol which can provide security 100% or near about it. Many researchers are working on it and they are searching for the best protocol which can provide security as much as possible. WiMaX is the recent technology in the Wi-Fi security. It also has some deficiencies.

References

Journal Papers:

- [1] Wireless security: an overview by Robert J. Boncella. Washburn University ZZbonc@washburn.bdu.
- [2] White paper: WLAN security Today: wireless more secure than wired by Siemens Enterprise Communications.
- [3] Sara Nasre Wireless Lan Security Research Paper IT 6823 Information Security Instructor: Dr. Andy Ju An Wang Spring 2004.
- [4] Security Issues on Converged Wi-Fi & WiMAX Networks by Prof. Anand Nayyar, Lecturer, P.G. Department of Computer Science, K. L. S. D College Ludhiana ,anand_nayyar@yahoo.co.in .

Theses:

- [5] *Wireless network security?* Author:-Paul Asadoorian, GCIA, GCIH. Contributions by Larry Pesce, GCIA , GAWN PaulDotCom.
- [6] *Securing Wi-Fi network* (10 steps of diy security) by Rakesh M Goyal and Ankur Goyal
- [7] *Establishing wireless robust security networks: a guide to IEEE 802.11i* by Sheila Frankel Bernard Eydt Les Owens Karen Scarfone.
- [8] *Wireless LAN security today and tomorrow* By Sangram Gayal And Dr. S. A. Vetha Manickam .
- [9] *Introduction to WI-FI network security* by Bradley Mitchell, About.com.
- [10] *The state of WI-FI security* by WI-FI Alliance.
- [11] *WI-FI security –WEP, WPA and WPA2* by Guillaume Lehembre.
- [12] *Wireless network security 802.11, Bluetooth and handheld devices* by Tom Karygiannis, Les Owens.
- [13] *WEP, WPA, WPA2 and home security* by Jared Howe.

Application Of Process Activity Mapping For Waste Reduction A Case Study In Foundry Industry

Mr. Girish. C. Pude¹, Prof. G. R. Naik², Dr. P. G. Naik³

^{1,2}Department of Production Engineering, KIT'S College of Engineering, Kolhapur, Shivaji University, Kolhapur (India)

³Chhatrapati Shahu Institute of Business Education & Research, Kolhapur

Abstract: Value Stream mapping technique involves flowcharting the steps, activities, material flows, communications, and other process elements that are involved with a process or transformation. In this respect, Value stream mapping helps an organization to identify the non-value-adding elements in a targeted process and brings a product or a group of products that use the same resources through the main flows, from raw material to the arms of customers. Process activity mapping is a Value stream mapping (VSM) tool has its origins in industrial engineering. The technique is known by a number of names although process analysis is the most common. Current effects of wastes on processes are observed. Process activity mapping steps are studied and wastes are identified. Different flow layouts are studied and best layout is selected that reduces the transport time. In this paper a case study in xyz foundry is carried out with current state map and future state map after following the different steps starting from the detailed time study of mapping process from raw material to final product. The steps including waste elimination techniques with conversion of existing waste into standard wastes are described. Statistical charts are prepared for the identification of bottleneck product is presented.

Keywords: Current State map, Future State Map, Process activity mapping, Value Stream Mapping, VSM tool, Waste Reduction.

I. INTRODUCTION

Process is a series of actions, changes or functions that bring about an end result. (American Heritage, 1978). A Process is defined as one or more tasks that transform a set of inputs into a specified set of outputs (goods or services) for another person i.e. customer or process via a combination of people, procedures, and tools. (Wesner, Hiatt, Trimble, 1994). The sequence of things [procedures] done to produce an output. A task is just one individual step in the process (Flanigann, Scott, 1995). Process mapping is the detailed mapping of the real process. Process maps: (a) Bring clarity to complex processes (b) Highlight non-value adding activities and (c) Start the process of thinking about improvements.

There are five stages to the process activity mapping [1]:

- (1) The study of the flow of processes;
- (2) The identification of waste;
- (3) A consideration of whether the process can be rearranged in a more efficient sequence;
- (4) A consideration of a better flow pattern, involving different flow layout or transport routing; and
- (5) A consideration of whether everything that is being done at each stage is really necessary and what would happen if superfluous tasks were removed.

A case study in xyz foundry for waste reduction with process activity mapping is carried out to identify the process flow and non value adding activities. The study is focused only on foundry production line (Line 1 namely KOYO production line) which contributes to 98% of the castings of the foundry. The Value stream mapping tool is used to analyze both the flow of materials and the flow of information. The line is semi-automated and the casting goes through various processes drums, conveyors, sand separation, sand preparation, knock out, degating, shot blasting etc. Mapping the value stream activities of other products would give similar results with small differences in numerical values as some castings are simple, requires hot/cold box process while some are complex requiring cold/hot box process. However the bottleneck product Unterlage is selected for the study. This study is a case study applied for the foundry processes which are continuous processes as production lines are semi automated and the analysis uses the process activity mapping tool. This tool is selected with questionnaires, interviews and brainstorming sessions with manager and operators on the shop floor [2]. Marketing, production planning & control dept., pattern shop has helped us to give past record values of foundry production line for e.g. marketing dept. has given monthly customer order, production planning dept. has given the total foundry layout, pattern shop has given classification of simple and complicated jobs depending upon their work experience etc.

II. Case Study

2.1 Objectives

1. To use the Value Stream Mapping tool in identifying, quantifying and minimizing major wastes in a foundry production line.
2. To quantify by rank the seven wastes of lean within the foundry line.
3. To formulate practical means of reducing the identified major wastes.
4. The aim is to reduce lead time for process improvement.

2.2 Methodology

Step 1 Identify the bottleneck product for case study.

Unterlage 2884 and the product family is the bottleneck product corresponding to customer Wagner Trident (Fig.1).

Step 2 Causes for bottleneck

2.1 The product selected for the study is Unterlage corresponding to customer Wagner Trident as this product family possesses more casting weight up to 12.5kg which is greater than any product or product family on the production line. See fig.1 showing the graph of customer versus average casting weights in kg.

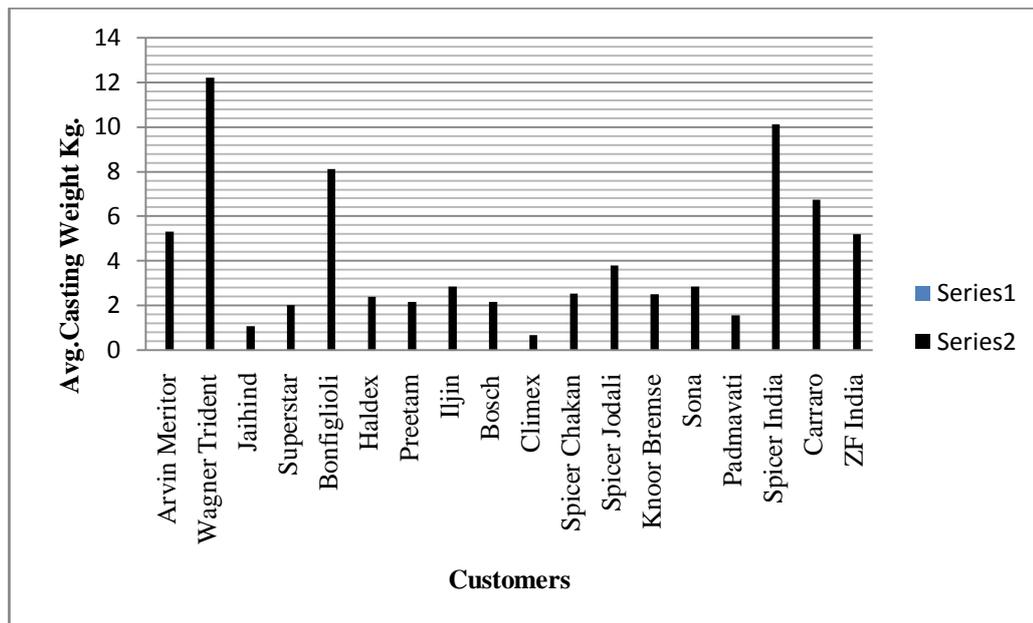


Fig.1 Bottleneck Product

2.2 Further the no. of cavities, pouring time and mould cycle times for Unterlage are greater than any other product or product family on production line1.

Step 3 Selection of the value stream mapping tool.

3.1 The Value Stream Analysis Tool (VALSAT) [1] is done to select the proper tool.

The value stream to be reviewed is first identified and through a series of preliminary interviews with managers the wastes are found out and ranked based on the weight age to a particular waste. The Value Stream analysis is done on the basis of correlation matrix of the seven wastes and the appropriate tools to use for analysis.

Table 1: The Seven Value Stream Mapping Tools [1]

Wastes / Structure	Mapping Tools						
	Process Activity Mapping	Supply Chain Response Matrix	Production Variety Funnel	Quality Filter Mapping	Demand Amplification Mapping	Decision Point Analysis	Physical Structure (a) Volume (b) Value
Overproduction	L	M		L	M	M	
Time Waiting	H	H	L		M	M	
Transport	H						L
Inappropriate Processing	H		M	L		L	
Unnecessary Inventory	M	H	M		H	M	L
Unnecessary Motion	H	L					
Product Defects	L			H			
Overall Structure	L	L	M	L	H	M	H
Origin of Tool	Industrial Engineering	Time compression/ Logistics	Operations Management	New Tool	Systems Dynamics	Efficient Consumer Response Logistics	New Tool

Notes: H = High correlation and usefulness
 M = Medium correlation and usefulness
 L = Low correlation and usefulness

According to VALSAT [1] three tools namely Process activity mapping, Quality filter mapping and Production variety funnel have shown greater effectiveness to reduce the waste in foundry. However only Process activity mapping is considered for analysis in this paper for case study.

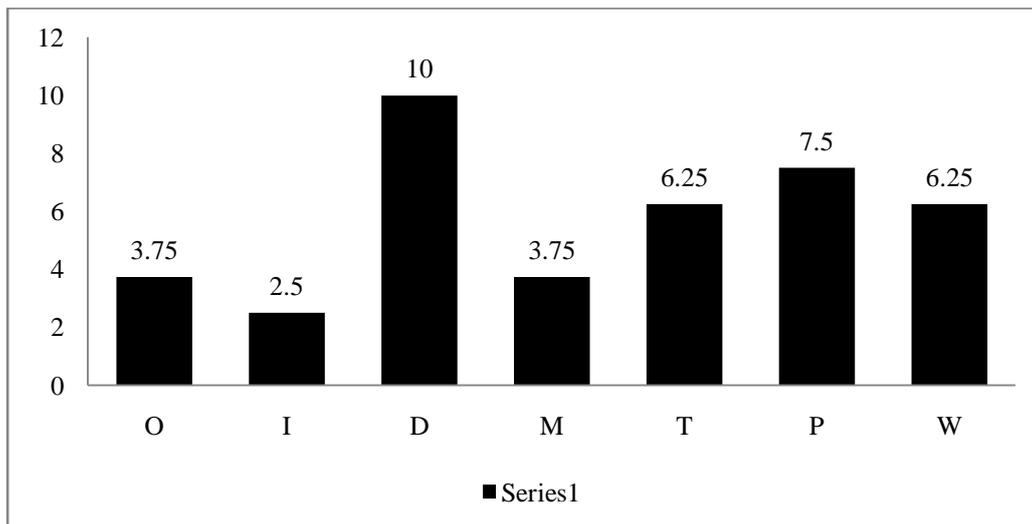
Step 4 Case details

The time study for all the foundry operations was carried out for 45 days by using stopwatch as a recording

technique. The data is collected for the *bottleneck product per machine per shift per component*. The statistical bar charts are drawn to reveal the product details and analyze the problem.

Step 5 Study of the manufacturing process flow

The xyz foundry production line is semi automatic dedicated to flow the product. The process flow is shown in Fig. 2. The process starts from raw material along with component drawing, sand preparation step by step. The



O=Overproduction, I=Inventory, D=Defects, M=motion, T=Transportation, P=Inappropriate Processing, W=Waiting
 Fig.3Waste Ranking Bar Chart.

Step 7 waste ranking methods

Based on the VALSAT analysis and the time study analysis the waste ranking algorithm is employed and it shows the wastes that really need to be minimized as shown in fig.3.

Step 8 Identification of Value Stream

The value stream is identified and it comprises of all the activities that are Value added and Non Value added. Identification of value stream will give the detailed information about the product flow, wastes that occur

Step 9 Mapping the Process Activities

Process activity mapping involving the preliminary analysis of the process [1] followed by the detailed recording of all the items required in each process is done for bottleneck product UNTERLAGE *per machine per shift per component* on KOYO automated line. As can be seen from this foundry example each step (1-38) has categorized in terms of variety of activity types (operation, transport, inspection and storage). The machine or area used for each of these activities is recorded, together with the distance moved, time taken and number of people involved. These are shown by darker red bold letters as shown in fig 4.

The total distance moved, time taken and people involved can be calculated and recorded [1]. The completed diagram (Figure 4) can then be used as the basis for further analysis and subsequent improvement. This is achieved through the use of techniques such as the 5W1H [1] (asking: *Why* does an activity occur? *Who* does it? *On which* machine? *Where?* *When?* and *How?*). The basis of this approach is therefore used to eliminate activities that are unnecessary, simplify others, combine yet others and seeks sequence changes that will reduce waste. Various contingent improvement approaches can be mapped similarly before the best approach is selected for implementation.

2.3 Current State Map

The time study is carried out by the stopwatch for 45 days on the shop floor of the foundry. We have gone through each and every process on the production line 1 by recording the travelling distances of men, materials, time taken by each activity, number of operators and workers right from raw material to final product dispatch. We

Waste reduction is possible by using different lean techniques such as creating supermarket, Kaizen, JIT, FIFO etc. approach which can be conducted through discussions with shop floor employees.

within the value stream, step by step activities with cycle times, distances etc. It gives systematic procedure to tackle the improvement opportunities with lean techniques to satisfy the customer demand with increase in productivity.

observed raw material inventory, in-process and finished good inventory. WIP is more in between the degating-shot blasting-inspection-dispatch area. Layout was improper and due to that mixing up of parts, low productivity, poor housekeeping and unnecessary transportation of parts was there in fettling section.

Current state map is prepared keeping in view of the lean manufacturing principles. A few assumptions are also made for preparation of current state map. From past sales data at the industry under study, it is known that maximum demand of Unterlage 2884 may reach up to 3300 per month. The current state map captures information at a particular instance, which may vary from shift to shift. For the sake of analysis, the shift and operator-wise variation (which may be there) is not considered. Effective numbers of working days are 26 per month, number of shifts per day is three and working hours per shift are eight. Available working time per day is 86400 seconds. As per [3] Takt time can be calculated as

$$\text{Takt time} = \frac{\text{Available working time per day (seconds)}}{\text{customer demand per day (units)}} = \frac{86400}{(3300/26)} = 681 \text{ seconds.}$$

From current state map, value added time as a % of total time in plant = $\frac{113497.54}{(15 \text{ Day} \times 24 \text{ hrs/day} \times 60 \text{ min/hr})} = 0.62 \%$.

2.4 Future State Map

For future state map the following areas in which wastes like WIP, Inappropriate processing, bad layout etc. are identified and presented separately in corresponding process. After that the wastes are converted into standard wastes and techniques of waste elimination are described in latter part. First the processes in which the identified wastes exists are, a) Pattern Making, b) Core Making, c) Moulding, d) Shot Blasting, e) Fettling

Sr.No	STEP	FLOW	MACHINE	DIST (M)	TIME (MIN)	PEOPLE	OPERATION	T	I	S	D	COMMENTS
								R	A	N	S	
								P	O	R	T	
								T	I	S	D	
1	RAW MATERIAL	S	RESERV-OIR	-	-	-	O	T	I	S	D	PIG IRON,MANGANESE,COPPER ETC
2	COMPONENT DRAWING	O	DRAWN BY ENGINEER	-	1440	-	O	T	I	S	D	1 DAY FOR SIMPLE 3 DAY FOR COMPLICATED (1 SHIFT/DAY)
3	MOULD PATTERN DESIGN MAKING & METHODING	O	-	-	12000	1	O	T	I	S	D	COMPLICATED SINGLE PIECE PER PLATE (25 DAYS). FOR COPE & DRAG IT REQUIRES TOTAL 50 DAYS.(1 SHIFT/DAY)
4	CORE PATTERN MAKING	O	-	-	-	-	O	T	I	S	D	FOUNDRY OUTSOURCE S CORE PATTERNS
5	RAW MATERIAL TRANSFER BY HOIST CRANE TO FURNACE	T	-	15.6	2	1+2	O	T	I	S	D	1 OPERATOR+2 WORKERS FROM STORAGE TO BASIN & BASIN TO FURNACE.
6	MELTING TIME	O	INDUCTION FURNACE	-	30	-	O	T	I	S	D	
7	POURING INTO LADLE	O	-	-	2	1+2	O	T	I	S	D	1OPERATOR + 2WORKERS

8	MOVING LADLE FOR POURING UPTO TANDISH FLASK	T	-	16.5	4	1+2	O	T	I	S	D	TANDISH FLASK IS DESIGNED FOR POURING INTO MOULDS (1OPERATOR +2WORKERS)
9	MOULD SAND PREPARATION(WAIT FOR MIX)	D	SAND MIXER TUMBLER	-	2	1+1	O	T	I	S	D	BINDER,ADDITIVES AND NEW SAND IS ADDED (1 OPERATOR+1WORKER)
Sr.No	STEP	FLOW	MACHINE	DIST (M)	TIME (MIN)	PEOPLE	OPERATION	T	I	S	D	COMMENTS
10	MOULD SAND TRANSFER BY CONVEYOR TO HIGH PRESSURE MOULDING LINE	T	-	21	0.3	-	O	T	I	S	D	
11	MOULD MAKING	O	HIGH PRESSURE MOLDING (KOYO)	-	0.25	2+1	O	T	I	S	D	NO. OF MOULDS=4(WITH CORE) AND=5(NON CORE)(2 OPERATOR+1 WORKER)

12	CORE SAND POURING TO CORE MAKING MACHINE	T	-	-	1	2	O	T	I	S	D	MANUAL POURING, RESIN BINDER SAND IS USED.
13	CORE SAND PREPARATION (WAIT FOR MIX)	D	SAND MIXER	-	0.25	-	O	T	I	S	D	SAND BOX TIME
14	CORE MAKING	O	CORE SHOOTER	-	1.5	1	O	T	I	S	D	SHOOTING PRESSURE 3-5 BAR (1 OPERATOR)
15	CORE HARDNESS TESTING	I	SCRATH HARDNESS TESTER	-	0.167	1	O	T	I	S	D	
16	POURING MOL TEN MET AL FROM TAN DISH TO MOL D	O	TANDISH FLASK	-	0.2	1	O	T	I	S	D	1 OPERATOR
17	POURED MOL D TO KNO CKOUT BY CON VEY OR	T	-	66	60	-	O	T	I	S	D	

18	KNOCK OUT	O	VIBRATOR	-	0.334	-	O	T	I	S	D	
Sr.No	STEP	FLOW	MACHINE	DIST (M)	TIME (MIN)	PEOPLE	OPERATION	TRANSPORT	INSPECTION	STORAGE		COMMENTS
19	FROM KNOCK OUT TO HARDNESS TESTING BY TROLLYS	T	FORK LIFT	32	5	2	O	T	I	S	D	
20	HARDNESS TESTING FOR CASTING	I	BRINELL HARDNESS TESTER	-	120	1	O	T	I	S	D	REQUIRES ONE HOUR FOR COOLING AND ANOTHER FOR TESTING
21	FROM HARDNESS TO CUTTING	T	-	16	3	2	O	T	I	S	D	
22	CUTTING THE CASTING	O	CUTTING GRINDER	-	5	1	O	T	I	S	D	CUTTING FOR MICROSTRUCTURE TESTING.
23	FROM CUTTING TO MICRO TESTING	T	-	80	3	2	O	T	I	S	D	MANUAL TRANSPORT

24	MICROSTRUCTURE TESTING FOR CASTING	I	ELECTRON MICROSCOPE	-	30	3	O	T	I	S	D	INCLUDES POLISHING & TESTING
25	FROM KNOCK OUT TO DEGATINING THROUGH DRUM & CONVEYOR	T	-	26.5	7	-	O	T	I	S	D	
26	DEGATINING	O	ONLINE WEDGE CUTTER	-	0.17	1+1	O	T	I	S	D	RUNNER, RISER SEPARATION FROM CASTING (1OPERATOR+1 WORKER)
Sr.No	STEP	FLOW	MACHINE	DIST (M)	TIME (MIN)	PEOPLE	OPERATION	T	I	S	D	COMMENTS
27	FROM DEGATINING TO SHOT BLASTING BY CONVEYOR	T	-	9.1	3	-	O	T	I	S	D	
28	SHOT BLASTING	O	SHOT BLASTING MACHINE	-	5	-	O	T	I	S	D	

29	FROM SHOT BLASTING TO FETTLING BY CONVEYOR	T	-	19	3	1	O	T	I	S	D	INCLUDES ONLINE PRIMARY INSPECTION (1 INSPECTOR)
30	FETTLING & POINT GRINDING	O	GRINDER	-	15	1+1	O	T	I	S	D	POINT GRINDER AND GRINDING WHEEL (1 OPERATOR+ 1 WORKER)
31	FROM FETTLING TO FINAL INSPECTION	T	MANUAL TRANSFER	4	0.5	1	O	T	I	S	D	
32	FINAL INSPECTION	I	-	-	3	1	O	T	I	S	D	MANUAL INSPECTION (1 INSPECTOR)
33	OIL DIP	O	OIL TUB	-	0.083	1+1	O	T	I	S	D	CARRIED OUT TO PREVENT RUSTING (1 OPERATOR)
34	AWAIT TRUCK	D	-	-	60	-	O	T	I	S	D	
35	AWAIT FILLING THE TROLLEY	D	-	-	2	2	O	T	I	S	D	

36	FRO M TRO LLY TO TRU CK BY FOR K LIFT	T	-	85	3	1	O	T	I	S	D	
Sr.No	STEP	F L OW	MACHINE	DIST (M)	TIME (MIN)	PEOPLE	O P E R A T I O N	T R A N S P O R T	I N S P E C T	S T O R E	D E L A Y	COMMENTS
37	WAI T TO FILL FUL L LOA D	D	LORRY	-	60	2	O	T	I	S	D	FILLING LOAD AS PER CUSTOMER ORDER
38	AWA IT SHIP MEN T	D	LORRY	-	420	1	O	T	I	S	D	DISPATCH
	TOT AL		38 STEPS	390.7	14291. 754	47	14	1 3	4	1	6	
	OPE RAT ORS				13497. 537	13						
	% VAL UE ADD ING				94.44 %	27.66%						

Fig.4 Process Activity Mapping

Identification of Wastes in the above foundry processes:

I) Pattern Making

1. Patterns must be finished, edges, burrs should be removed.

II) Core making

1. Three cavity core boxes for component 1 were not enough to complete production requirement.
2. Core Box of 2 cavities & core for component 2 was solid creating gasses problem in castings and not enough to complete production requirement.
3. One cavity core box for component 3was not enough to complete the production requirement.
4. Core weight is more for Slip Yoke 1180 which is 0.370 kg.

III) Moulding

1. Spillover of dust in turn affecting the environment in mould sand plant.
2. Moulding sand hopper was damaged so that sand leakage was more in the mould making section.

IV) Shot Blasting

1. Due to m/c breakdown of 3-4 days WIP is more in this section.

V) Fettling

1. Mixing up of different parts due to bad layout.
2. Poor housekeeping leads to low productivity.
3. High transportation of parts.

Classification of wastes into standard forms and techniques of elimination adopted:-

Table.3 Techniques of wastes elimination.

Operations	Waste	Type	Technique of Elimination
Pattern Making	1.Patterns must be finished, edges, burrs should be removed	Inappropriate Processing	Abrasive grinding brush introduced.
Core Making	1.Three cavity core boxes for component 1 were not enough	Machine Error	Design and Process Review.10 Cavity Core Box is made.
Operations	Waste	Type	Technique of Elimination
	2. Two cavity core box for component 2 was solid creating gasses problem in castings.	Machine Error	Design and Process Review.6 Cavity Core Box is made.
	3. One cavity core box for component 3 was not enough.	Machine Error	Design and Process Review.2 Cavity Core Box is made.
Moulding	1. Spillover of dust in environment.	Machine Error	Design and Process Review. Containers were provided for collecting the dust.
	2. Moulding sand hopper was damaged causing sand leakage.	Machine Error	Design and Process Review. Hopper welding with proper direction was done.
Shot Blasting	1. M/c breakdown caused more WIP.	Waiting	Total productive Maintenance (TPM) or Preventive Maintenance.
Fettling	1. Mix up of different parts due to poor layout.	Waiting	Part Specific self concept was adopted.5S was used.
	2. Low productivity due to bad poor housekeeping.	Waiting	Part Specific self, 5S was used.
	3. High transportation of parts	Transportation	Cellular Manufacturing or Group technology.

From future state map, value added time as a % of total time in plant = $\{(13497.54 / (11 \text{ Day} \times 24 \text{ hrs/day} \times 60 \text{ min/hr})\} = 0.85 \%$. The future state map indicates average 23% waste reduction in the critical areas of unnecessary inventory, transportation and waiting.

The concept of supermarket came into picture to reduce the in process inventory or WIP. A “supermarket” is nothing but a buffer or storage area located at the end of the production process for products that are ready to be shipped (Rother and Shook, 1999). On the other hand, producing directly to shipping means that only the units that are ready to be shipped are produced. Currently xyz foundry produces all the automotive Spheroidal Graphite Iron products (S.G Iron) and sends them to a holding area where they are stored with other products waiting to be shipped. However, this is done based on a push system, and components can wait a long time in this area before being shipped. The introduction of supermarkets is

necessary only at the finishing end where large amounts of inventory exist between different workstations. In addition to the shipping supermarket recommended after final inspection, three additional supermarkets are needed to create a continuous flow. One before the Shot blasting, one before the fettling process and one before final inspection as flow displayed in Fig. 6. First in between degating and shot blasting ,2nd in between shot blasting and fettling,3rd in between fettling and inspection & 4th after the final inspection before the dispatch. Instead of Kanban pull, manual push system is employed to fill the supermarket to their required capacity. Kaizen continuous improvement programmes were carried out at pattern making, core making, mould making and fettling departments with Total Productive Maintenance (TPM) and 5s as explained above in table 3. First In First out (FIFO) concept was employed before degating, shot blasting and fettling.

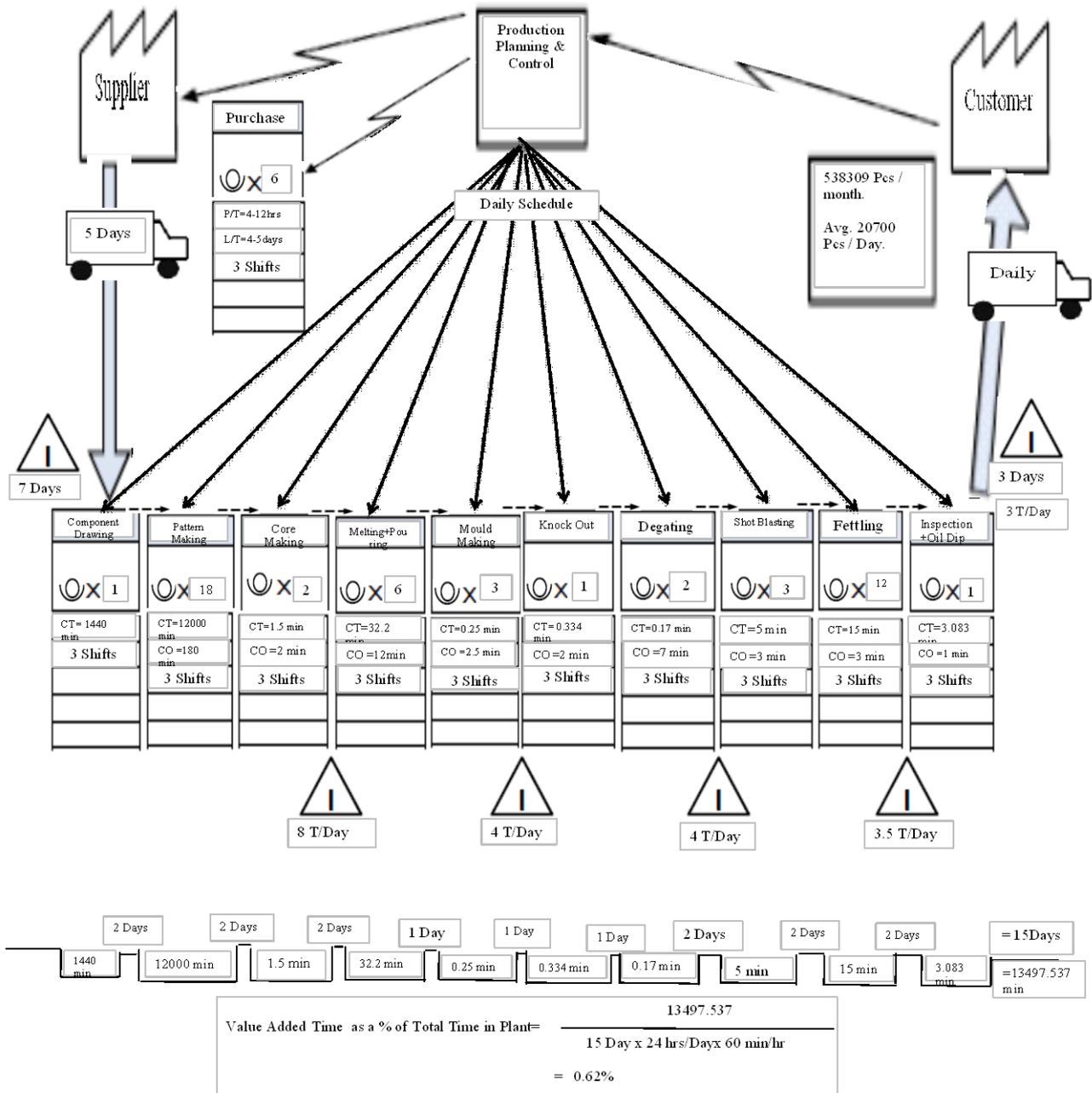


Fig 5. Current State Map

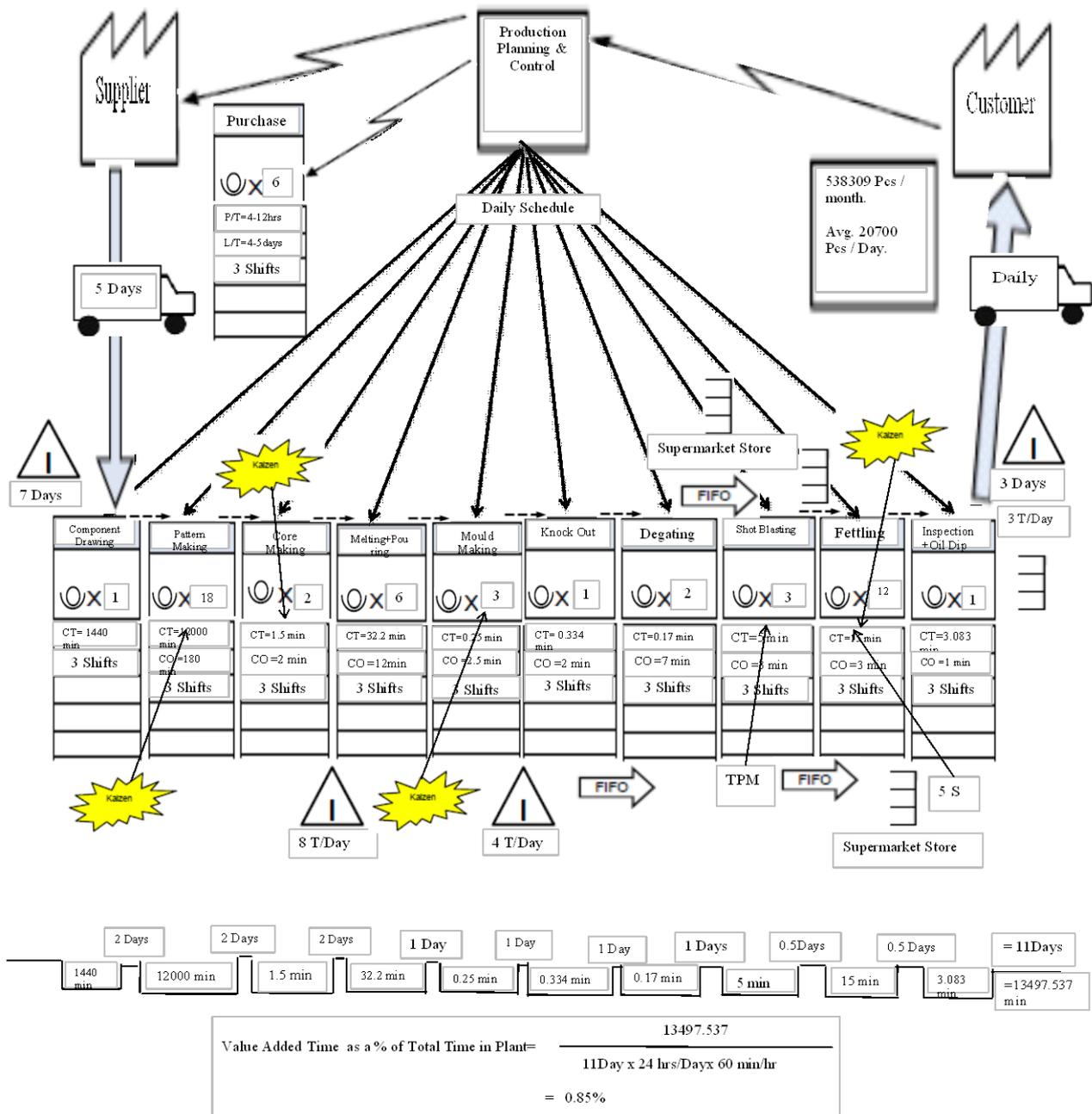


Fig 6. Future State Map

III. Conclusion

Value stream mapping tool can be used effectively in any kind of sectors as it is a world class manufacturing tool. The analysed study is case study in foundry industry. The prime objective is to carry out Process Activity Mapping for waste reduction. In this study bottleneck product was identified. Further Value Stream Analysis Tool (VALSAT) was employed for identification of wastes in a process. The significance of each type of error was studied and waste ranking was carried out. Current state map is plotted to assess current status. Waste elimination techniques are presented and future state map is also preferred for improvement. The results of study shows 23% waste reduction in the areas of unnecessary inventory, transportation and waiting. It is however to be noted that there is a significant cost to complete any required changes

but the increased throughput against takt time will pay back for investment.

The authors are currently working on this project in a foundry situated in Maharashtra (India). This study is carried out on foundry production Line 1(KOYO), which produces 98% of total castings in foundry, therefore study is focused on 1st production line. A similar analysis on other production line 2(DISA) may show similar trend of results. Heavy complex castings are produced on second line contributing to only 2% of total castings. Besides authors are also working on Production Variety Funnel and Quality Filter mapping tools for process improvements. Value stream mapping tools can be effectively employed to reduce wastes and to improve the process.

REFERENCES

- [1] Hines P, Rich N. "The seven value stream mapping tools" *Int J Operations and production Management* 17 46–64 1997.
- [2] William M. Goriwondo, Samson Mhlanga, Alphonse Marecha "Use of the value stream mapping tool for waste reduction in manufacturing-Case study for bread manufacturing in Zimbabwe" *International Conference on Industrial Engineering and Operations Management Kuala Lumpur, Malaysia* 2011 236-241.
- [3] Bhim Singh, Suresh K. Garg, and Surrender K. Sharma "Value stream mapping: literature review and implications for Indian industry" *Int J Advanced Manufacturing Technology* 53 2011 799–809.
- [4] Ulf K. Teichgraber, Maximilian de Bucourt "Applying value stream mapping techniques to eliminate non-value-added waste for the procurement of endovascular stents" *European Journal of Radiology* 81 2010 e47– e52.
- [5] Fawaz A. Abdulmalek, Jayant Rajgopal "Analyzing the benefits of lean manufacturing and value stream mapping via simulation: A process sector case study" *Int. J. Production Economics* 107 2006 223–236.
- [6] Process Mapping & Procedure Development 2007 Consultas Pty Ltd.

Real-time Actuation of Cylindrical Manipulator model in OpenGL based on Hand gestures recognized using Open CV

Subash chandar¹, Willson Amalraj², Gnana sambandam³

Department of Electrical and Computer Engineering,
National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077.

Abstract: This paper proposes a method of actuating robotic cylindrical manipulator model in OpenGL based on hand gestures recognized using OpenCV (Open Source Computer Vision Library) in VC++. A model of cylindrical manipulator is developed using OpenGL (Open Graphics library) in VC++, which is actuated for the recognized hand gestures. The fingers and their orientation with respect to a reference base plane are recognized for different gestures using OpenCV. The actuator in each joint of cylindrical manipulator is selected by specific hand gesture. The movement of the actuator is controlled by orientation of the hand with respect to base plane. The hand gesture recognition and cylindrical manipulator module are developed using Windows Application Program Interface (WINAPI), standard C++ libraries, OpenCV libraries and OpenGL libraries in visual studio 2010. Real time simulation of the prototype model developed in VC++ is demonstrated with successful gesture recognition and actuation of manipulator.

Keywords: Hand gestures, cylindrical manipulator, Open Source Computer Vision Library (OpenCV), Open Graphics Library (OpenGL), Windows Application Program Interface (WINAPI), Visual C++.

I. INTRODUCTION

Human interaction with computers is constantly evolving in order to increase the efficiency and effectiveness by which interactive tasks are performed. Nowadays, devices such as keyboards, mouse, joysticks or gamepads are used for human-computer interaction (HCI). In the recent years a growing interest in methods of natural actions recognition in human-to-human communication such as speak and gesture in a more effective way is seen [1]-[2]. Hand Gestures which has become popular in recent years can be used to develop a Human computer interaction device [3]-[6]. Hand gestures recognition by Computer will provide a more natural human-computer interface, allowing people to point, or rotate a CAD model in a virtual environment, by rotating their hands. Gesture recognition may even be useful to control household appliances [7].

Hand gestures can be recognized by vision-based or data glove based methods. Vision-based hand gesture recognition methods are based on the quality of the captured images [8]-[10]. Because of the limitations of the optical sensors to lighting conditions and cluttered backgrounds, vision based approaches are not efficient in detecting and tracking the hands robustly, which largely affects the performance of hand gesture recognition [11]. Data glove such as Cyber glove can be used to enable a more robust hand gesture recognition which effectively captures the hand gesture and motion using embedded

sensors [12]-[13]. Unlike optical sensors, Cyber gloves are more reliable and are not affected by lighting conditions or cluttered backgrounds. However, it requires the user to wear a data glove and sometimes requires calibration which is inconvenient for the user. Moreover the cost of data glove is much higher when compared to conventional webcam used for vision-based approach of hand gesture recognition. This paper deals with Computer vision based gesture recognition system [14].

The gestures are classified into two categories such as static and dynamic. A static gesture is represented by a single image corresponding to a particular hand configuration and pose. A dynamic gesture is a moving gesture, represented by a sequence of images. In this project dynamic gesture recognition is used to identify the finger counts in the recognized hand, which is done by processing the video acquired from a webcam [15]-[17]. The finger counts obtained for specific hand gestures are used to select the appropriate actuator in a cylindrical manipulator model developed in OpenGL. The orientation angle obtained from the hand gesture recognition system is used to control the movements of actuators in the manipulator model. The selection and control of actuators in manipulator is done effectively in real-time without time delay.

The rest of the paper is organized as follows. In section II, a brief overview of OpenCV is presented. Section III gives an introduction to OpenGL. The proposed architecture for image processing and actuation of robotic cylindrical manipulator is discussed in detail in Section IV. In section V, the results of the proposed system are discussed. Section VI concludes the paper and the future works are presented in section VII.

II. OPEN CV

The OpenCV (Open Source Computer Vision Library) is a cross-platform library of programming functions developed by Intel for real-time computer vision. It is free for use under the open source BSD license. Its main focus is on the real-time image processing. The library will make use of Intel's Integrated Performance Primitives, if available in the system which has optimized routines to accelerate itself. OpenCV aids advanced vision research by providing not only open but also performance optimized code for basic vision infrastructure. The library was originally written in C which makes OpenCV portable to some specific platforms such as digital signal processors. However since version 2.0, OpenCV includes both its traditional C interface as well as a new C++ interface. The number of lines of code for vision functionality is reduced in OpenCV 2.0. It also reduces common programming errors such as memory leaks through automatic data allocation and de-allocation that may arise when using OpenCV in C. wrappers in other languages to C++ is not

developed as opposed to C code. OpenCV libraries can be accessed by Visual C++ in windows operating systems. However the Base Classes from DirectShow SDK is required to access webcams on Windows with DirectX Media SDK prior to 6.0. Facial recognition system, gesture recognition, human computer interaction (HCI), motion tracking, etc. are a few applications of OpenCV.

III. OPEN GL

OpenGL (Open Graphics Library) is a cross-language, multi-platform application program interface developed by Silicon Graphics Inc., for developing and simulating applications, that produce 2D and 3D computer graphics. OpenGL accepts primitives such as points, lines and polygons, and converts them into pixels via a graphics pipeline known as the OpenGL state machine. The interface can be used to draw complex three-dimensional scenes from simple primitives. OpenGL presents a single, uniform interface for different 3D accelerator hardware and supports full feature set, using software emulation if necessary, for all implementations. OpenGL is a low-level, procedural API, requiring the programmer to have a good knowledge of the steps required to render a scene. Hence it provides certain amount of freedom to implement novel rendering algorithms in contrast with descriptive APIs, where a programmer only needs to describe a scene and can let the library manage the details of rendering it. Several libraries are built beside OpenGL to compensate for features not available in OpenGL. Libraries such as GLU, GLUT, SDL and Allegro have been developed for aiding OpenGL for rudimentary cross-platform windowing and mouse functionality. Simple graphical user interface functionality can be found in libraries like GLUT or FLTK. OpenGL is widely used in CAD, virtual reality, scientific visualization, information visualization, flight simulation, and video games.

IV. PROPOSED SYSTEM ARCHITECTURE

The architecture has two different modules, Image processing and cylindrical manipulator module working with synchronization. The flow diagram of the proposed algorithm is shown in Fig.1. The detailed working principle of each module in the algorithm is discussed in this section.

A. Image processing Module

The Image processing module process the image acquired to recognize the gesture. The entire module is developed in VC++ using OpenCV libraries and WINAPI (windows application programming interface). The code is developed in 32 bit, windows7 OS, hence 32 bit version of OpenCV and WINAPI dynamic link libraries. This module consists of different stages with specific functions to identify the hand and to recognize the gesture. The functionality of each stage is discussed in detail in the following sections.

1) Image Acquisition

The image of the hand is acquired through webcam, it can be either integrated or standalone. Webcams can be directly accessed using OpenCV libraries or DirectX libraries present in windows SDK (Software development kit). The webcam is configured such that the frame rate is 30 and the resolution is fixed to be 480 x 640 (VGA). The

capturing resolution can be increased depending on the webcam available, which in turn will increase the processing time of the algorithm.

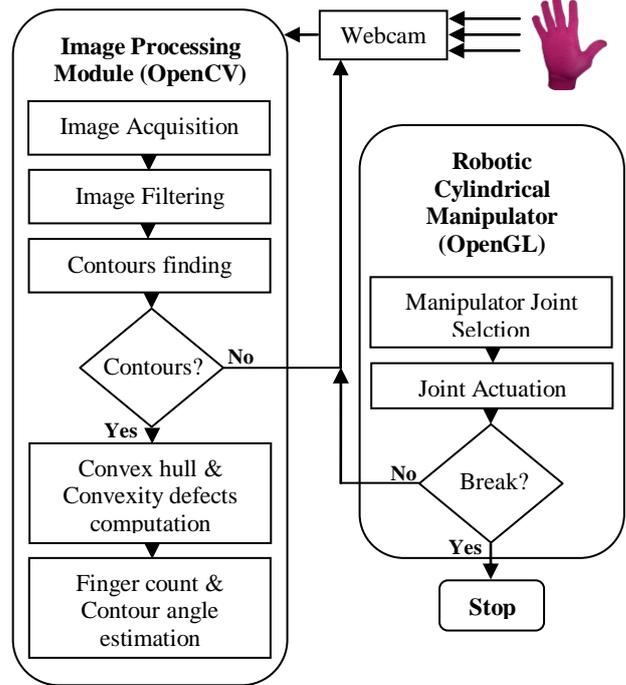


Fig.1 Flow diagram of proposed system

2) Image filtering

The captured image is filtered to extract the pixels retaining the information pertaining to the hand. Attempt was done to recognize the bare hand, but hard coding it, will not be useful, as the complexion differs for different person. Moreover if a range is fixed for filtering the image, it will lead to other objects in surrounding to appear in the filtered image. Hence the natural choice for easy hand detection will be a glove. The program is hard coded to recognize the hand with pink or blue glove.

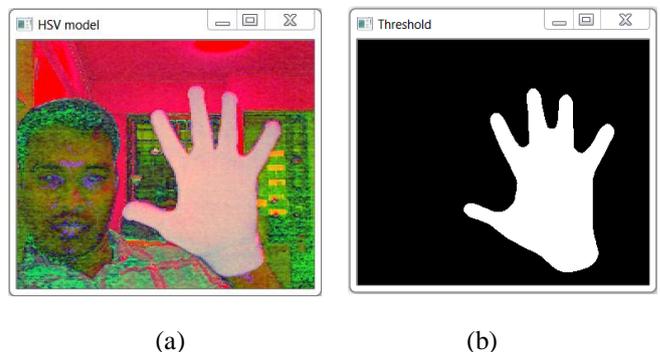


Fig.2 Image captured using webcam (a) HSV format & (b) Final filtered image using CvThreshold

The image captured by webcam in BGR format (3 channels) is flipped to match with the same direction as that of gesture movement in front of webcam. The BGR format is converted to HSV format using cvCvtColor() in OpenCV as shown in Fig.2 (a), which is then filtered to remove the backgrounds except hand, using cvInRange(). cvInRange() compares the source image with the corresponding value in the lower and upper arguments. If the value in the source image is greater than or equal to the value in lower

argument and also less than the value in upper argument, then the corresponding value in destination will be set to 1 otherwise, the value in will be set to 0. Hence the image filtered after `cvInRange()` will be of binary format (1 channel). The noise in the image is then removed by Median filter using `cvSmooth()`. Finally the required pixels with dominant features are separated from the other pixels using `cvThreshold()` as shown in Fig.2 (b).

3) Contours, convexity Hull and Convexity Defects Computation

The final filtered image is used to compute the contours with `cvFindContours()` function. The function `cvFindContours()` computes contours from binary images or images with edge pixels. Contours are represented in OpenCV by sequences in which every entry in the sequence encodes information about the location of the next point on the curve. Once the contour is obtained the next logical step is to comprehend the shape of the processed image.

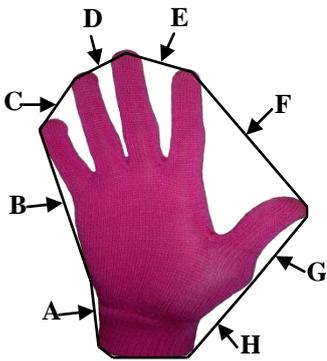


Fig.3 Convex hull and convexity defects of a hand

The convex hull of the contour obtained from recognized hand is pictured as a dark line around the hand in Fig.3, and the regions labeled A through H are each “defects” relative to that hull. Convexity defects characterize not only hand but also its state. Hence the number of fingers can be easily estimated for different gestures.

The number of fingers is used to select the specific actuator in the manipulator. The direction of manipulator actuation has to be controlled by some other controlling parameters. Orientation of the hand contour is chosen as the controlling parameter. The detail of the hand contour formed is summarized with a bounding ellipse using `cvEllipseBox()`. The angle of the bounded ellipse is used to control the direction of actuator movement.

B. Robotic Cylindrical Manipulator Module

A cylindrical manipulator is designed using OpenGL and WINAPI in VC++. The manipulators consist of three joints, a prismatic joint for translating the arm vertically, a revolute joint with a vertical axis and another prismatic joint orthogonal to the revolute joint axis. The each segment of robotic manipulator is designed by a scaled cube using `glutSolidCube()` function in OpenGL as show in Fig.4. If a cube is drawn in OpenGL the center of the cube will be at local coordinate of system. Hence, if the cube is made to rotate, it will do so about its center rather than the base. This is avoided by moving the local coordinate system to one edge of the cube by calling appropriate modeling transformation `glTranslate()`. The same procedure is

repeated for all the segments in the model to orient with respect to base cube which serves the platform.

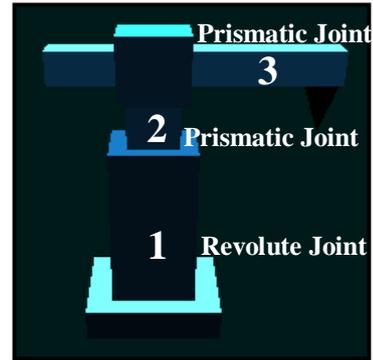


Fig.4 Robotic cylindrical manipulator in OpenGL

The rotation and translational movement is achieved by establishing pivot points for each segments, using `glRotate()` and `glTranslatef()` respectively. In OpenGL the transformations are done by matrix multiplication, hence the position and orientation of the coordinate system for each segment must be saved and restored at appropriate moments. Failure to do so will result in loss of degree of freedom for other segments. This is overcome by using `glPushMatrix()` and `glPopMatrix()` functions in OpenGL. The function `glPushMatrix()`, copies the current matrix and adds the copy to the top of the stack, and `glPopMatrix()`, discards the top matrix on the stack. Whenever a segment is made to move, the contents of the current window will be changed and it has to be redisplayed. It is done by declaring all the routines needed to redraw the scene in the display callback function, `glutDisplayFunc()`.

The material properties of the objects in the scene like, the ambient, diffuse, and specular colors, the shininess, and the color of any emitted light are also declared using `glMaterialfv()` in display callback function. The function `glLightfv()` is used to set the proprieties such as color, position, and direction of light sources in the scene which are declared in a separate one time call function and it remains unchanged.

1) Manipulator Joint Selection

The cylindrical manipulator has three joints capable of translational or rotational movement. Each joint is selected by a specific gesture recognized by image processing module. The gestures and their corresponding joint are depicted in Table I.

TABLE I. HAND GESTURES AND DIRECTION OF ACTUTATION

Hand gestures	Finger Count	Direction of actuation	Actuator
 Contour angle > 135°	1	Counter clockwise	Revolute Joint

 Contour angle < 135°			
 Contour angle > 135°	2	Up	Prismatic Joint 1
 Contour angle < 135°		Down	
 Contour angle > 135°	3	Backward	Prismatic Joint 2
 Contour angle < 135°		Forward	

The primary feature extracted from the gesture for joint selection is finger count. Finger count is computed from convexity defects of recognized hand. The count one corresponds to revolute joint, two for prismatic joint with vertical movement, and three for prismatic joint with horizontal movement as in Table I.

2) Manipulator Joint Actuation

The joint actuation is achieved by computing the hand contour angle computed in image processing module. The angle of hand contour is the angle subtended by major axis of ellipse which bounds the hand contour, with respect to the base plane (x-axis). The angle subtended by hand contour on quadrant-II is considered for actuation. As all the joints in manipulator have two degrees of freedom, it can be made to move in two different directions. The angle subtended is divided into two regions, 135° to 180° for one direction and 90° to 135° degrees for other direction. The joints and their corresponding movements with respect to two different set of angles are shown in Table I.

V. RESULTS AND DISCUSSION

The image processing module (OpenCV) is run as a thread while the robotic cylindrical actuator (OpenGL) module is the main function. Once the program starts execution, the OpenGL module will appear and will wait for response from OpenCV module.

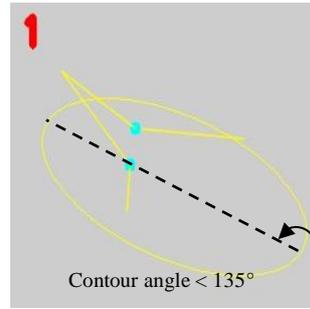


Fig.5 Revolute joint selection and actuation

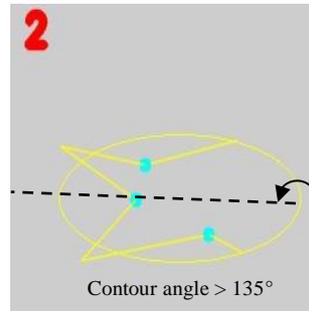


Fig.6 Prismatic joint 1 selection and actuation

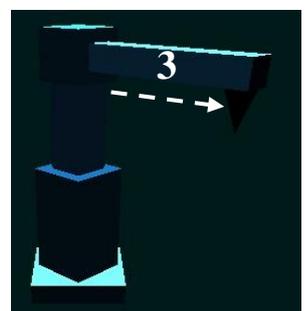
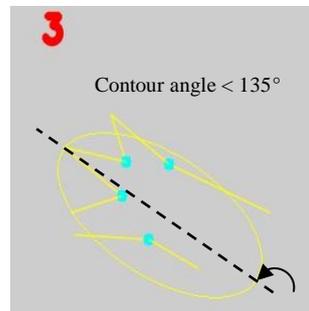


Fig.7 Prismatic joint 2 selection and actuation

When a hand with pink glove is exposed to webcam, contour is formed and the computation of finger count and contour angle will take place in OpenCV thread. Depending on the finger count appropriate joint is selected whose actuation depends on the contour angle. The finger count of hand contour, selected joint and direction of joint actuation are shown in Fig.5 to Fig.7. As the gesture is recognized dynamically the joint selection and actuation takes place without any considerable delay. As OpenCV and OpenGL modules use their specific dynamic linked libraries present in their respective software packages a redistributable package is created using Visual studio 2010, which can be run in other PCs without VC++.

VI. CONCLUSION

The hand gestures are recognized in real-time using OpenCV module developed in VC++. The dynamic gestures are recognized with good accuracy without any time lag. The recognized gesture is able to actuate the robotic cylindrical manipulator designed in OpenGL in real-time. The proposed method of hand gesture recognition and actuation is successful in recognizing and actuating in real-time.

VII. FUTURE WORK

A communication module has been developed for interface with the real robotic cylindrical manipulator in VC++. A serial port interface with a microcontroller is designed to provide communication between software modules running in a PC and the robotic manipulator. The communication module has been tested for different gestures and it is able to generate good result. Eventually the future work will be integrating the OpenCV, OpenGL and communication module with a real robotic cylindrical manipulator. The actuation of real robotic cylindrical actuator will be evaluated.

REFERENCES

- [1] O'Hagan, R. & Zelinsky, A. (1997) "Finger Track – A Robust and Real-Time Gesture Interface". Australian Joint Conference on AI, Perth.
- [2] Agrawal, T. and Chaudhuri, S., 2003. Gesture Recognition Using Position and Appearance Features, ICIP pp. 109–112.
- [3] Von Hardenberg, C. and Berard, F., "Bare-hand human-computer interaction", in [PUI '01: Proceedings of the 2001 workshop on Perceptive user interfaces], 1{8, ACM, New York, NY, USA (2001).
- [4] Malassiotis, S. and Srinivasan, M. G., "Real-time hand posture recognition using range data", *Image Vision Comput.* 26(7), 1027-1037 (2008).
- [5] McAllister, G., McKenna, S. J., and Ricketts, I. W., "Hand tracking for behavior understanding", *Image Vision Comput.* 20(12), 827-840 (2002).
- [6] Ionescu, B., Coquin, D., Lambert, P., and Buzuloiu, V., "Dynamic hand gesture recognition using the skeleton of the hand", *EURASIP J. Appl. Signal Process.* 2005(1), 2101-2109 (2005).
- [7] William T. Freeman and Michal Roth "Orientation histograms of hand gesture Recognition", IEEE Intl. Workshop. on Automatic Face and Gesture Recognition, Zurich, June, 1995.
- [8] Ueda, E., 2003. A Hand Pose Estimation for Vision-Based Human Interfaces, *IEEE Transactions on Industrial Electronics*, Vol. 50, No. 4, pp. 676–684.
- [9] Davis, J. and Shah, M., 1994. Visual Gesture Recognition, *Vision, Image and Signal Processing.* 141(2), pp. 101–106.
- [10] Pavlovic, V.I., Sharma, R. & Huang, T.S. (1997), "Visual Interpretation of Hand Gestures for Human- Computer Interaction: A Review. In *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 19 (7) 677-695.
- [11] Jingling Meng, Junsong Yuan, "Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction", 8th International Conference on Information, Communications and Signal Processing (ICICS) 2011.
- [12] Subash chandar, Vignesh Kumar and Willson Amalraj, "Real Time Data Acquisition and Actuation via Network for Different Hand gestures using Cyber glove", *International Journal of Engineering Research and Development*, ISSN: 2278-067X, Volume 2, Issue 1 (July 2012), PP. 50-54.
- [13] Jiangqin, W., wen, G., yibo, S., wei, L., and bo, P., "A simple sign language recognition system based on data glove", *Signal Processing Proceedings, 1998. ICSP 98. 1998 Fourth International Conference on*, 2, 1257{1260 vol.2 (1998).
- [14] G.R. Bradski, V. Pisarevsky, "Intel's Computer Vision Library", *Proc of IEEE Conference on Computer Vision and Pattern Recognition, CVPR00*, v. 2: 796-797, 2000.
- [15] Oka, K., Sato, Y., and Koike, H., "Real-time fingertip tracking and gesture recognition", *Computer Graphics and Applications, IEEE* 22, 64-71 (Nov/Dec 2002).
- [16] Malima, A. et al. 2006. "A fast algorithm for vision-based hand gesture recognition for robot control", 14th IEEE Int. Conf. on Signal Processing and Communications Applications, Antalya, Turkey.
- [17] Brethes, L. et al. 2004. "Face Tracking and hand gesture recognition for human robot interaction", *Int. Conf. on Robotics and Automation*, Vol. 2, pp. 1901–1906 New Orleans.

Static and Dynamic Analysis of Thin-Walled Cyclic Shells

S. N. Krivoshapko, DSc¹, Christian A. Bock Hyeng, Ph.D.²

¹(Professor, Peoples' Friendship University of Russia, 6, Miklukho-Maklaya Str, Moscow, 117198, Russia)

²(North Carolina Agricultural and Technical State University, 1601 E. Market St., Greensboro, NC 27411, USA)

ABSTRACT: The principal achievements of science and engineering in the sphere of static, vibrational, and buckling analysis of thin-walled structures and buildings in the shape of cyclic surfaces with circular generators are summarized in this review paper. These shells are useful as fragments of pipelines, spiral chambers of refrigerating units, as well as in spiral chambers of turbines in hydroelectric power stations, in high pressure units, in public and commercial buildings, for example, as coverings of stadiums, and so on. This review paper contains 62 references, and these are practically all original sources dealing with static and dynamic analysis of thin-walled cyclic shells.

Keywords: Canal surfaces of Joachimsthal, Cyclic surfaces with circular generators, Dupin's cyclides, epitrochoidal shell, spiral chamber, static and dynamic analysis, strength, tubular shell

I. Introduction

To date great progress has been achieved in the strength analysis of thin elastic shells. Thin-walled shell structures combining lightness with considerable strength find widespread use in modern engineering and building. Many have noted the rapid progress of the practice and the theory of the application of thin shells and thin-walled shell structures in the last 10 years. But shells used in constructions belong to limited classes of surfaces. An accurate tendency in world practice is the application of spatial structures of arbitrary form, giving expressive architectural images and solving functional problems. Today we have a new generation of young architects and engineers who have also shown interest in designing wide-span spatial coverings. This process amplifies occurrence of new materials, such as fibrous reinforced polymeric composites, which can be used in covers due to its ability to curve. One family of generating curves in cyclic surfaces is represented by circles of constant or variable radii that considerably reduces the cost price and simplifies the process of manufacturing thin shells in the form of these surfaces without a decrease in operational possibilities. However, until the middle of 20th century the analytical method for calculation of cyclic shells was replaced by the approached calculation concerning simple systems on which it was possible to dismember a cyclic structure. Design research on models was widely applied. Engineers, mechanics, and architects, using only the rough methods of analysis, basically spent on the basis of intuitive reasons, have created

many interesting structures and buildings in the form of cyclic surfaces.

In connection with inquiries of practice spanning 60 years to the present time, attention of many scientists attracts the problem of the determination of the stress-strain state of shells of complex geometry including cyclic shells;

therefore, in the presented review considerable attention will be paid to this problem. Such cyclic surfaces as surfaces of revolution, circular helical surfaces, and tubular surfaces are the most known and widely used. In this review, surfaces of revolution won't be considered, as hundreds of publications are devoted to them, and these surfaces are usually allocated in a separate class.

II. The review of works on research of strength of tubular shells

A paper [1] and a dissertation [2] were the first works on strength analysis of tubular shells. A. Bantlin [1] has experimentally established that a pipe with the curvilinear axis under bending is more flexible than a pipe with a rectilinear axis of the same cross-section [3]. The effect of T. Karman consisting of flattening of the cross-section of a pipe under bending raises flexibility of curvilinear pipes. I.V. Stasenko [4] determined that distortion of a contour of a cross-section is accompanied by the occurrence of the circular bending moments and normal stresses which can be several times more than that of nominal stresses of bending calculated without the effect of flattening. Let b be the radius of the cross-section of a pipe, a the radius of a centerline, and h the thickness of the pipe wall. When the parameter b^2/ah is small in comparison with unity then the influence of flattening does not matter. Using a trigonometric series for reception of quantitative results it is necessary to take the larger number of members, as it has more value than the parameter b^2/ah . T. Karman [3] considered a problem about a pipe of round cross-section using a principle of a minimum of potential energy in the form of the Rayleigh – Ritz. G. Lorentz has applied a principle of the least work of Al. Castigliano. R. Clark and E. Reissner [5] for the first time having applied a method of asymptotic integration have shown that G. Lorentz's theoretical results are less exact.

In the monograph [3], the short review of results of strength research of tubular shells with the references [1, 2, 5-8] is presented. Without repeating the above-stated review, we will specify only that in a work [6] the scheme of a semi-analytical method of final elements of calculation of curvilinear tubular shells with arbitrary spatial axis is investigated. Here approximation of fields of displacements in a shell in the form of the final sum of Fourier's series for one variable and oblique-polynomial approximation of type of final element method for another variable was applied. In a work [8], a thin-walled pipe is modeled by toroid and described by the equations of semi-momentless shell theories.

A series of works [9 – 12] devoted to strength analysis of tubular shells with application of approximate momentless theory was published by V.N. Ivanov. He showed that a system of three equilibrium equations for a momentless shell referred to lines of principal curvatures may be reduced to one resolving differential equation in

two ways. Introduction of function of stresses $\varphi(\alpha, \beta)$ [9, 10] is the first way. An exception of two inner forces of momentless state and reception of the resolving equation concerning the one remaining force is the idea of the second method [9, 13]. The analysis of the resolving equation with function $\varphi(\alpha, \beta)$ has shown that a method of separation of variables cannot be applied. Exceptions make torus and circular cylindrical shells. The tubular shell of arch type with directing flat curve is considered in a work [11]. For solution of the problem, a difference method of straight lines was used. Derivatives on the coordinate coinciding with the cross-section of a pipe are replaced by difference relations. Thus, resolving equations of momentless theory are reduced to systems of ordinary differential equations with variable coefficients. The tubular shell with a line of centers in the shape of evolute of the circle subjected to dead weight has been calculated with the help of momentless theory in a work [10]. As a result, resolving the difference equation became the Airu equation and its decision was received in Bessel functions.

The new scheme of discretization (a method of curvilinear grids) which is a generalization of a finite difference method is offered in [14]. It completely excludes an error of approximation of the covariant derivative of functions of rigid displacements. An analysis of stress-strain state of shells in the form of tubular surface with the flat sinusoidal centerline used for the connection of two parallel pipelines of identical radius is fulfilled in the monograph [3] with application of a method of curvilinear grid. Gulyaev and his colleagues put finite difference netting of dimensions 17×17 and 21×21 on the chosen area for control of convergence of calculations. The analysis of results of calculations has shown that the stress-strain state of shells under uniform internal pressure appreciably depends on geometrical parameters, and the maximum values of displacements, membrane forces, and bending moments take place in the joining of a tubular shell to parallel cylindrical shells.

Results of a series of experiments carried out on a breadboard model of the heat exchange tubes for the determination of fatigue failure in the dryout zone over frequent heat changes are reviewed in a work [15]. The technique of carrying out of modeling tests in welded helical pipes of the heat exchanger is described. A.S. Pystogov [16] has presented governing formulas for definition of tangent stresses and deformations of helical springs of tension-and-compression with coils of tubular section from thick-walled pipes. Simmonds James G. [17] has derived geometrically nonlinear parities defining deformations, turns and changes of curvatures of coordinate lines, the equations of compatibility of deformations, the equilibrium equations, and Hooke's law with reference to analysis of a helicoidal circular cylinder (a tubular helical surface, see Fig. 1) with an inextensible contour subjected to axial tension and torsion. A method of separation of variables for analysis of the stress-strain state of tubular helical shells was realized for a two-dimensional case in papers [18, 19] where properties of this shell depending on values of its parameters are also investigated. It should be noted that stresses and moments under small face pitch do not surpass corresponding levels in toroidal shells [19].

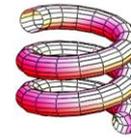


Fig. 1

Attempts at making analytical decision of concerned problems with the stress-strain state of tubular shells meet great difficulties of a calculating nature. The various rough methods are applied to the decision of the resolving differential equations, for example, momentless shell theory and simplifications.

The additional information on static behavior of tubular shells can be found in [20-28] and in paper [21], research in which a modified finite difference method has been allowed to replace the system of differential equations in private derivatives by a system of algebraic equations. The extensive bibliography of works devoted to analysis of pipes with a curvilinear axis resulted in R. Clark and E. Reissner's manuscript [5].

Elastic equilibrium of tubular shells was considered widely enough but the dynamic behavior of these shells are examined only in [3, 29-32]. In the monograph [3], it is noticed that today's research on vibrations of pipelines and their connecting sites have a special significance due to the growth of capacities of machines and velocities of movement of working liquids and gases. Frequencies and forms of natural vibrations are the basic dynamic characteristics of any engineering structure. Calculation of frequencies and forms of natural vibrations of a connecting element of two parallel circular pipelines is made in [3]. The rigid fixation was taken on the two opposite circular sides. On the allocated area of calculation, the finite difference net by dimension 17×17 was put and it provided sufficient accuracy of calculations. Analyzing the received results, the authors [3] notice that the curvature of a centerline appreciably has an influence on character of change of frequencies of its natural vibrations and on distribution of forms corresponding to them. Sokolov [30] has defined frequencies of natural vibrations of the curvilinear pipeline with a flowing liquid on a basis of momentless theory of thin-walled shells of "mean bending." Vibrations of pipes with a flat centerline with taking into account flowing viscous incompressible liquid are also considered by Ya. F. Kayuk in [31]. The account of effect of flowing liquid in a helical pipe was investigated in a work [33]. Influence of filling of a helical pipe by a saturated porous medium was examined in [34]. Only one work [35] devoted to experimental study of helical pipes has been found.

III. The review of works on research of epitrochoidal shell strength

Geometry and working with methods of analysis of epitrochoid shells using momentless theory have been considered for the first time in a master's thesis [36] by Mahmud Hussein Suleiman. Geometrical characteristics of epitrochoid shell give an opportunity to reduce a system of equations of momentless theory of shells due to a function of stress $\varphi(\alpha, \beta)$ to one equation

$$\frac{2}{\tilde{R}(\alpha)} \frac{\partial}{\partial \alpha} \left[\tilde{R}^2(\alpha) \frac{\partial \varphi}{\partial \alpha} \right] + \frac{1}{\omega(\alpha, \beta) f} \frac{\partial}{\partial \beta} \left[\frac{\omega^2(\alpha, \beta)}{f} \frac{\partial \varphi}{\partial \beta} \right] = G(\alpha, \beta), \quad \omega(\alpha, \beta) = \xi(\alpha) - f^2(\beta);$$

$\xi(\alpha) = [2\sigma^2(\alpha) - \tilde{R}(\alpha)] / \tilde{R}^3(\alpha)$; $\tilde{R}(\alpha) = R(\alpha)/a = 1 + \mu \cos \alpha$; $\sigma = \sqrt{\tilde{R}^2 + \tilde{R}'^2}$; where G is a function of load. Dependence of the function ω on parameters α, β does not generally allow for application of a method of separation of variables for the resolving equation. The analysis of dependence of function $\xi(\alpha)$ from parameter μ has shown that for each value μ it is possible to define an interval of change of coordinate parameter α in which $\xi(\alpha)$ will be a slowly changing function and it is possible to replace its value with an average value of this function in the set interval. So, it gives an opportunity to use a method of separation of variables [37]. Calculation of fragments of the epitrochoid shell with $\alpha = 0, \alpha_0 = \pi/4, \beta = \pm 1$ and with parameters $\mu = 0.1; 0.3; 0.5$ has been carried out (Fig. 2). The shell is loaded by own weight q . Conditions of supporting of the fragment provides equality to zero of tangent forces at the edges $\beta = \pm 1$ and $\alpha_0 = \pi/4$. Normal forces N_α are equal to zero on the edge $\alpha = 0$.

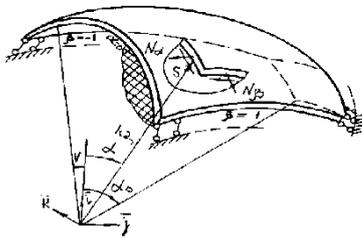


Fig. 2

The function of stresses was accepted in the form

$$\varphi(\alpha, \beta) = qa^3 \left[\varphi_0(\alpha, \beta) + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{mn} \Phi_m(\alpha) F_n(\beta) \right],$$

$$\Phi_m(\alpha) = \sin(m\alpha); \quad F_n(\beta) = G_{2n-1}^c(\beta) - C_n G_{2n+1}^c(\beta);$$

$G_n^c(\beta)$ are orthogonal polynomials of Gegenbauer which are functions of the resolving equation. Similar calculations have been made for a compartment symmetrical on coordinate $\alpha (-\pi/4 \leq \alpha \leq \pi/4)$. Results of calculation are presented in [38] in graphical representation.

Research of stress-strain state of shells in the shape of epitrochoidal surfaces have been continued in the dissertation [39]. It was shown that epitrochoidal surfaces belong to a class of canal of surfaces of Joachimsthal. Using approaches to analysis of epitrochoidal shells offered in [36, 37], Gil Oulbe Mathieu has carried out a strength analysis of the part of epitrochoidal shell, limited by edges $\alpha = \pm \alpha_0, \alpha_0 = \pi/4, \beta = 0$ and $\beta = 1$. The line of centers lies in a horizontal plane (Fig. 3). Boundary conditions provide equality of tangent forces on the contour. Research of the convergence process of calculation was realized with various number of members of a series. Researchers have shown that in the main part of a shell, accuracy of calculation is provided in the 10th approach. For the support zone $\beta = 0$, convergence has been reached in the 18th approach [40]. The results of calculation are presented in graphical representation.

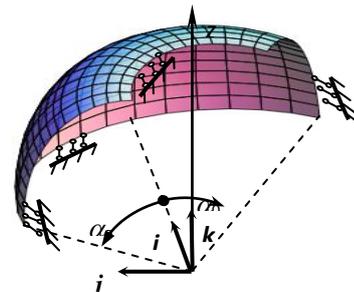


Fig. 3

IV. The review of works on research of strength of canal surfaces of Joachimsthal

Research of geometry and stress-strain state of shells in the form of canal surfaces of Joachimsthal are discussed in a dissertational work [43]. The equation of canal surfaces of Joachimsthal with arbitrary directrix in lines of principal curvatures, coefficients of the first and second fundamental forms in the theory of surfaces and principal curvatures have been received in [44]. The investigation of methods of analysis of shells has shown that analytical methods of analysis are not applicable for analysis of shells of complex forms even if one uses momentless theory of shells. Different numerical methods are used for moment analysis of shells of complex forms. For analysis of thin shells having the shape of the canal surfaces of Joachimsthal, a finite difference energy method has been chosen. The choice has been made on the basis of the analysis of possibilities of three methods: a difference method of netting, a finite element method, and a finite difference energy method. Systems of the equations describing stress-strain state of shells in the form of Joachimsthal's surfaces are used in a difference method of netting directly. The order of derivatives can reach the 4th – 8th order. Necessity to satisfy to all boundary conditions of a shell structure extremely complicates the realization of this problem, especially near curvilinear edges, in a zone of openings and on free shell edges. A finite element method and a finite difference energy method are based on the minimization of functional of total energy of deformations, the derivatives in which not above the 2nd order. The methods allow satisfying only the kinematic boundary conditions, realization of which usually does not represent any difficulties. The FEM is convenient for calculation of shells with openings and with curvilinear contour of the shell not coinciding with a coordinate net. Complex

contours and openings are easily approximated by selection of special final elements.

This problem becomes complicated if one uses a finite difference energy method, but it is quite realistic. A FEM has entered into settlement practice in the 2nd half of the 20th century and has won great popularity and it is widely applied for analysis of thin-walled spatial structures. Program complexes realizing calculations of various shell structures are created. However, the equations of middle surfaces in FEM are used usually only for creation of a net of knots of finite elements on the middle surface of the shell meant for analysis. The geometrical characteristics reflecting features of internal and external geometry of a surface usually are not used in FEM. It brings additional errors in results of calculation of thin shells of complex geometry. For analysis of shells of complex geometry by finite difference energy method, an algorithm considering geometrical characteristics of middle surface of shells had been developed [45-48]. The program complex using FORTRAN algorithmic language has been developed in Peoples' Friendship University of Russia for realization of the developed algorithm. The library of curves and surfaces is connected to the program module calculating values of geometrical characteristics of the middle surfaces of shells used for the determination of their stress-strain state. The program complex allows analyzing shells of canonical geometry and shells of complex geometry, including shells in the shape of canal surfaces of Joachimsthal. At debugging of the program complex, analyses of plates, spherical and cylindrical shells have been carried out. The results of calculation were compared to the known exact decisions. The calculations have shown good convergence at calculation of shell structures of cylindrical form on a net 8x8 and spherical form on a net 10x10 [49].

Having used a complex program for analysis of shells with the help of a finite difference energy method, N.Yu. Abbushi [43] carried out an analysis of a shell in the shape of Joachimsthal's canal surfaces with circular sinusoidal directrix $r_2(\alpha_1) = a[1 + \mu(1 + \cos k\alpha_1)]$ (Fig. 4). Shell calculation was fulfilled on the action of dead weight with $c = 0.707$; $\mu = 0.1$; $k = 8$; $f(\alpha_2) = \text{tg}\alpha_2$. The shell is closed along the coordinate α_1 ($0 \leq \alpha_1 \leq 2\pi$) but $0 \leq \alpha_2 \leq 2/3\pi$. The contour edge $\alpha_2 = 0$ is rigidly fixed and the upper edge $\alpha_2 = 2/3\pi$ is free. Values of internal forces and moments were received and resulted in the diagrams [12, 43]. An analysis of the closed shell in the shape of canal surfaces of Joachimsthal with oval directrix $r_2(\alpha_1) = \sqrt{l_1^2 \cos^2 \alpha_1 + l_2^2 \sin^2 \alpha_1}$ was also performed. The form of an oval curve depends on the relation of dimensional parameters $\lambda = l_1/l_2$ (Fig. 5). It was assumed that $a = 20$ m, $f(\alpha_2) = \tan \alpha_2$, that the material of the shell was homogeneous with $E = 2 \cdot 10^4$ MPa, and that the Poisson coefficient of the material was $\nu = 0.15$; a thickness of the shell was $h = 0.10$ m. The shell was loaded by the inner constant pressure $P_{in} = 10$ KH/m² (Fig. 6). Tables of values of the internal tangential normal forces and the bending moments are presented in [12, 43].

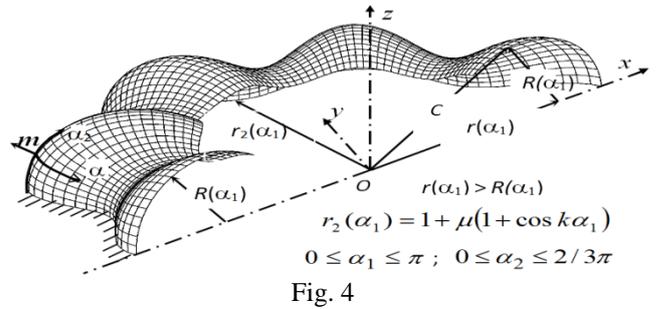


Fig. 4

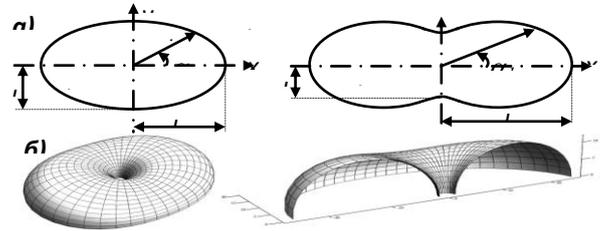


Fig. 5

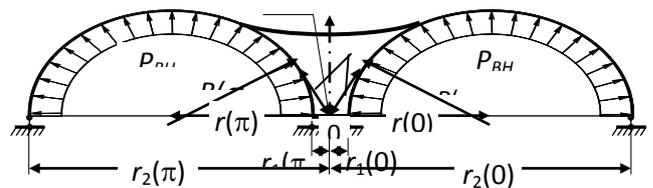


Fig. 6

V. The review of works on research of strength of shells in the shape of Dupin's cyclides

The fullest materials closely associated with definition of the stress-strain state of shells in the shape of Dupin's cyclides loaded by a uniform surface load are presented in [12, 50]. First, Krishna Reddi in his dissertation [51] and papers [52-54] presented numerical results of analysis of shells in the shape of Dupin's cyclides. He used a momentless shell theory. A momentless theory of shells in the shape of Dupin's cyclides of the fourth order of the first type (Fig. 7) was offered in [52]. An article [54] devoted to momentless theory of shells in the form of the third order cyclides of the second type, Fig. 8. Methods of analysis of cyclides presented by Krishna Reddi [51-54] are based on the representation of coefficients of the first fundamental form and principal radii as

$$A = \frac{f_3(\alpha)f_4(\beta)}{f(\alpha,\beta)}, B = \frac{f_5(\alpha)f_6(\beta)}{f(\alpha,\beta)}, R_1 = f_1(\alpha)f_2(\beta), R_2 = R_2(\alpha),$$

where

$$f_1(\alpha), f_2(\beta), f_3(\alpha), f_4(\beta), f_5(\alpha), f_6(\beta), f(\alpha,\beta)$$

are any functions. In this case, the system of equations of equilibrium is reduced to one resolving differential equation of the second order concerning the put function with private derivatives. This differential equation is solved in double trigonometric series for both types of Dupin's cyclides.

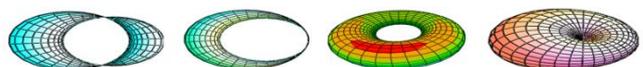


Fig. 7

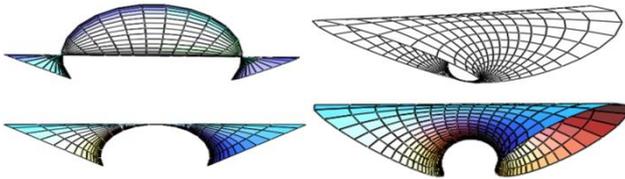


Fig. 8

Having changed coefficients f_i , S.A. Duheisat [55] has solved an additional example for a shell in the form of Dupin's cyclides with the help of momentless theories. For reception of two resolving differential equations of moment shell theory, Krishna Reddi [51] used the equations of a linear theory of shells in the complex form which have been received for the first time by V.V. Novozhilov.

Apparently, the works set forth above [50-56] limit the list of the works devoted to research of the stress-strain state of Dupin's cyclides. Dupin's cyclides of the fourth order of the third type, which are circular torus, are not examined in this review.

VI. Additional information on analysis of cyclic shells

The most well-known results devoted to analysis of cyclic shells subjected to static and dynamic actions are stated above in sections 1-4. In the present section, we shall note only several works devoted to some problems of cyclic shells which have not been mentioned elsewhere.

D.V. Vainberg and V.I. Gulyaev [57] have presented analysis of a spiral shell of the circular cross-section loaded by uniform external pressure and by a temperature field. The problem was solved numerically by a method of nets. L.S. Panasyuk [58] has shown the isometric nature of helical circular cylinder and torus. The method of complex limitations, having wide application in dynamic calculations in Western European countries, has been applied in [59] to definition of natural frequencies of spiral heat exchangers.

Analysis of spiral chambers of turbines of HYDROELECTRIC POWER STATION (Fig. 9) [60-62] possesses some specificity in comparison with a moment theory analysis of cyclic shells with generatrix circles of variable radius and with a flat line of centers constructed around the circular cylinder (Fig. 10), therefore strength analysis of shells of this type is not considered in this review.



Fig. 9

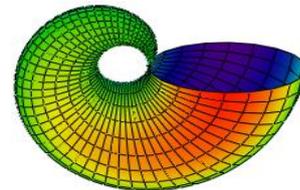


Fig. 10

Vibration of cyclic shells is examined in a monograph [3]. In particular, free vibration frequencies and corresponding mode shapes of the connecting channel for two cylindrical pipes with parallel axes and with different diameters (Fig. 11) are defined. It was established that the curvature of a centerline appreciably influences on character of change of nature frequencies and on the distribution of corresponding mode shapes. Special influence on dynamic behavior of shells renders the relation of initial and final radii of the generatrix circles. The curvature of a centerline leads to increase of value of the lowest frequency but reduction of parameter R_{in} / R_{fi} is accompanied by sharp decrease in its value. In the already mentioned monograph [3], frequencies and forms of own vibration of shells in the form of a normal cyclic surface with an elliptic centerline and with a generatrix circle of variable radius (Fig. 11) are investigated also. The deviation of an axial line from a circle leads to decrease in frequencies of own vibration both in shells with constant generatrix circles and in shells with variable radius of cross-sections. But the deviation of an axial line doesn't render essential influence on character of change of the shapes of own vibrations of these shells.

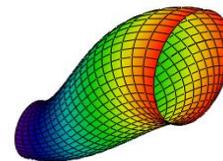


Fig. 11

VII. Conclusion

The materials contained in 62 monographs, dissertational works, scientific papers, proceedings of conferences, and listed in the review include almost all data on static and dynamic analysis of cyclic shells known in the world today.

As to engineering possibilities of application of cyclic shells, they have found wide application in curvilinear pipes (<http://www.ttb.com/>, Tulsa Tube Bending), basically in the form of torus, in spiral chambers of water-wheels, and in the circular translation shells used as spatial coverings of buildings.

Cyclic surfaces with a straight center line, or shells of revolution, are widely used in architecture. Sometimes architects specially underline the presence of generatrix circles in the design of buildings. As to cyclic structures of the general non-degenerate type, only single examples of their application are known. It is the authors' hope that literature resulting from the review and analysis of tendencies of development of strength analysis of cyclic shells will help architects, civil engineers, and mechanical engineers expand and diversify projected spatial structures using cyclic surfaces and will assist post-graduate students choose themes of scientific research.

References

- [1] Bantlin A. (1910), Formanderung und Beanspruchung Ausgleichsrohren, *Z. Ver. Deut. Ing.*, № 54, 43-49.
- [2] Wissler H. (1916), Festigkeitsberechnung von Ring flächenschalen, Technische Hochschule, *Diss.*, Zürich.
- [3] Gulyaev V.I., Bazhenov V.A., Gozulyak E.A., Gaydaychuk V.V. (1990), An Analysis of Shells of Complex Form, Kiev, Budivelnik, 192 c.
- [4] Stasenko I.V. (1980), An influence of initial departures on stress state of thin-walled curvilinear pipes, *Proc. MVTU*, № 332, 146-160.
- [5] Clark R., Reissner E. (1955), Bending of pipes with curvilinear axis, *Problemy mehaniki*, Moscow: IL, 125-149.
- [6] Muha I.S., Savula Y.G., Shinkarenko G.A. (1981), On analysis of tubular shells with arbitrary curvilinear axis, *Soprotivl. materialov i teoriya soor.*, Kiev, Is. 39, 71-74.
- [7] Noor A.K., Peters J.M. (1983), Recent advances in reduction methods for instability analysis of structures, *Computers and Structures*, vol. 16, No 1-4, 67-80.
- [8] Whatham J.F. (1981), Thin shell analysis of circular pipe bends, *Trans. Institution of Engineers*, Australia, vol. CE23, No 4, 234-245.
- [9] Ivanov V.N. (1971), An analysis of shells in the form of cyclic surfaces, *Diss. of PhD*, Moscow: UDN.
- [10] Ivanov V.N. An analysis of a tubular shell with a line of the centers in the form of evolvent of the circle subjected to dead load, *Sb. nauchn. rabot aspirantov ing. fakult.*, vol. 8, Moscow: UDN, 143-151.
- [11] Ivanov V.N. (1970), An analysis of tubular shells with the help of momentless theory, *Proektirovanie metal. konstr.*, vol. 7(27), Moscow: TzNIIS, 88-95.
- [12] Ivanov V.N., Krivoshapko S.N. (2010), Analytical Methods of Analysis of Shells of Non-canonical Form: Monograph, Moscow: RUDN, 540 p.
- [13] Rekach V.G. (1970), On a problem of integration of equations of momentless shell theory, *Tr. UDN*, vol. 48, № 6, Moscow: UDN.
- [14] Gozulyak E.A., Ermishev V.N., Zhadrassinov N.T. (1981), Convergence of a method of curvilinear nets in problems of a shell theory, *Soprotivl. materialov i teoriya soor.*, Kiev, Is. 39, 80-44.
- [15] Martin Ph., Dumas J.C., Girard J.P. (1986), Thermomechanical stresses in the dryout zone of slightly inclined helically coiled heat exchange tubes, *Boiler Dyn. and Contr. Nucl. Power Stat.*, Proc. 3rd Int. Conf., Harrogate, 21-25 Oct., 1985, London, 1986, 29-36.
- [16] Pystogov A.S. (1978), On analysis of screw springs with tubular cross-section, *Ural. PI, Sverdlovsk, 12 p.*, Dep. v NIInformtyazhmash, 26 apr. 1978, № 305.
- [17] Simmonds James G. (1984), General helicoidal shells undergoing large, one-dimensional strains or large inextensional deformations, *Int. J. Solids and Struct.*, 20, № 1, 13-30.
- [18] Gavelya S.P., Sysoev Y.A., Ryzhkova T.G. (1980), On the research of SSS of a tubular-and-helical shell, *Prikl. zadachi mat. fiziki*, vol. 1, Kiev, pp. 154-166, Dep. v VINITI 18 sept. 1980, № 4102-80Dep.
- [19] Gavelya S.P., Ryzhkova T.G. (1975), On the application of scheme of separation of variables for analysis of SSS of a tubular-and-helical shells, *Dopov. dokl. AN USSR*, № 7, 587-603.
- [20] Soboleva O.G. (1989), The bending of thin-walled pipe with curvilinear axis, *LISI*, 18 p., Dep. v VINITI 21.07.89, № 4894-B89.
- [21] Grigorenko Y.M., Gulyaev V.I., Gozulyak E.A., Ashuri K. (1983), Stress-strain state of tubular shells under action of uniform surface pressure, *Prikl. Mekhanika*, Kiev, № 8.
- [22] Svetlichnaya V.B. (2005), A strength analysis and rational design of curvilinear pipes made of reinforced plastics under static loading, *Diss. of PhD*, Volgograd, 163 p.
- [23] Kostovetskiy D.L. (1973), Strength of Pipe Lines of Energy Machines, 264 p.
- [24] Bagmutov V.P., Tyshkevich V.N., Svetlichnaya V.B. (2008), A Strength Analysis and Rational Design of Curvilinear Pipes Made of Reinforced Plastics: Monograph, Volgogrds: VGTU, 157 p.
- [25] Kruglyakova V.I. (1972), On analysis of thin-walled pipes with curvilinear axis, *Izv. AN SSSR, MTT*, № 6, 160-170.
- [26] Whathan J.F., Tompson J.J. (1979), The bending and pressuring of pipe with flanged tangents, *J. of Nucl. Eng. Desg.*, Vol. 54, № 1, 17-28.
- [27] Levyakov S.V., Pavshok V.N. (2008), Buckling analysis of flanged curvilinear pipes in pure bending, *Int. J. of Pressure and Piping*, Vol. 85, Iss. 5, May 2008, 306-312.
- [28] Vasudevaiah M., Patturaj R. (1994), Effect of torsion in a helical pipe flow, *Inter. J. of Mathematics and Mathematical Sciences*, Vol. 17, Iss. 3, 553-560.
- [29] Ashuri K. (1984), Natural frequencies of cyclic shells with a plane curvilinear line of the centers, *Soprotivl. materialov i teoriya soor.*, Kiev, Is. 44, 96-103.
- [30] Sokolov V.G. (1981), Natural vibrations of curvilinear pipelines, *Stroit. truboprovodov*, № 6, 25-26.
- [31] Kayuk Ya.F. Nonlinear vibrations of pipes with flowing viscous liquid, *Prikl. mehanika*, Kiev, №12, 89-97.
- [32] Sinitsin E.N., Chirkov V.P., Saharov V.Yu., Shmelyov D.N., Vlasov D.V. (2002), The vibrations of tube pencils in the flow of heat bearer, *The 2nd Conf. of Minatom of Russia "Methods and computer programs for strength analyses"*: Proc., Gelendzhik, 30 sept.-5 oct. 2002, Izd-vo GUP NIKIET, 2003, 56-64.
- [33] Yamamoto K., Aribowo A., Hayamizu Y., Hirose T., and Kawahara K. (2002), Visualization of the flow in a helical pipe, *Fluid Dynamics Research.*, Vol. 30, Iss. 4, April 2002, 251-267.
- [34] Nield D.A., Kuznetsov A.V. (2004), Forced convection in a helical pipe filled with a saturated porous medium, *Int. J. of Heat and Mass Transfer.*, Vol. 47, № 24, 5175-5180.
- [35] Shijie Liu, Artin Afacan, Hisham A. Nasr-El-Din, and Jacob H. Masliyan. (1994), An experimental study of pressure drop in helical pipes, *Proc. R. Soc. Lond. A.*, 444, pp. 307-316.
- [36] Mahmud Hussein M. Suleiman. (1992), The solution of problems of analysis of thin elastic shells in the form of epitrochoidal surfaces, *Diss. of PhD*, Moscow: RUDN, 150 p.
- [37] Ivanov V.N., Mahmud Hussein M. Suleiman (1992), An algorithm of calculation of epitrochoidal shell with the help of momentless shell theory, *Voprosy prochnosti prostr. sistem*: Proc. XXVIII scient. conf. of eng. faculty, Moscow: RUDN, 58-63.
- [38] Ivanov V.N., Mahmud Hussein M. Suleiman (1994), An analysis of an epitrochoidal shell with the help of momentless shell theory, *Raschet i proektir. grazhd. i gidrotehnich. soor.*: Mezhevuz. sb. nauchn. tr., vol. 3, Moscow: "Biokontrol", 21-26.
- [39] Gil Oulbe Mathieu (1997), An analysis of epitrochoidal shells in forces and in displacements, *Diss. Of PhD*, Moscow: RUDN, 134 p.
- [40] Ivanov V.N., Gil Oulbe Mathieu (1994), On a problem of geometry and design of shells in the form of canal surface of Joachimsthal, *Structural mechanics of engineering constructions and buildings*, Vol. 4, 68-75.
- [41] Gil Oulbe Mathieu (1996), An algorithm of calculation of epitrochoidal shell with the help of momentless shell theory in displacements, *Structural mechanics of engineering constructions and buildings*, Vol. 7, 18-21.

- [42] Ivanov V.N., Gil Oulbe Mathieu (1997), An example of determination of displacements of epitrochoidal shell subjected to own weight with the help of momentless shell theory, *Teoret. i experim. issled. prochnosti i zhostkosti elementov stroit. konstruk.*: Sb. tr., Moscow: MGSU, 82-86.
- [43] Nasr Yunes A. Abbushi (2002), Geometry, design, and research of stress-strain state of shells in the shape of canal surfaces of Joachimsthal, *Diss. Of PhD*, Moscow: RUDN, 149 p.
- [44] Ivanov V.N., Nasr Yunes A. Abbushi (1997), Geometrical research of canal surfaces of Joachimsthal, *Problemi teorii i praktiki v ing. issledovaniyah*: Proc. XXVIII scient. conf. of eng. faculty, Moscow: RUDN, 115-118.
- [45] Ivanov V.N., Nasr Yunes A. Abbushi (2000), An analysis of shells of complex geometry with the help of a finite difference energy method, *Structural mechanics of engineering constructions and buildings*, Vol. 9, 25-34.
- [46] Ivanov V.N., Nasr Yunes A. Abbushi (2000), The application of a finite difference energy method for an analysis of shells of complex geometry, *Tr. molodih uchonyh*, part 1, S.-Peterburg, 98-102.
- [47] Ivanov V.N., Nasr Yunes A. Abbushi (2001), An analysis of shells in the form of canal surfaces of Joachimsthal subjected to own weight with the help of a finite difference energy method, *Architecture of shells and strength analysis of thin-walled civil and machine-building structures of complex form*: Proc. of Intern. Scientific Conference, Moscow, June 4-8 2001, 121-126.
- [48] Ivanov V.N., Nasr Yunes A. Abbushi (2001), An analysis of shells in the form of canal surfaces of Joachimsthal by a finite difference energy method with the help of computer, *Spatial constructive systems of buildings and structures, methods of analysis, design and technology of erection*: Proc., Minsk, Oct. 10-12 2001, 76-83.
- [49] Nasr Yunes A. Abbushi (2001), Examples of analysis of civil structures by a finite difference energy method, *Structural mechanics of engineering constructions and buildings*, Vol. 10, 42-53.
- [50] Rekatch V.G. and Krivoshaplo S.N. (1988), Calculation of Shells of Complicated Geometry, Moscow, UDN, 177 p.
- [51] Krishna Reddi (1966), A strength analysis of shells in the form of Dupin's cyclides, *Diss of PhD*, Moscow, UDN, 157 p.
- [52] Krishna Reddi (1967), A momentless theory of shells in the form of Dupin's cyclides, *Issledovaniya po teorii soor.*, Vol.15, 90-99.
- [53] Krishna Reddi (1970), On equations of momentless theory of some types of shells, *Stroit. Mekhanika*, Vol. 48, Iss. 6, Moscow, UDN, 148-156.
- [54] Krishna Reddi (1967), Momentless theory of shells in the form of Dupin's cyclides of the third order, *Izv. vuzov, Stroitelstvo i arhitektura*, № 7, 47-55.
- [55] Salman al Duheisat (1987), On a problem of analysis of a momentless shell in the form of Dupin's cyclides, *Issled. po raschetu elementov prostranstv. sistem*, UDN, 106-110.
- [56] Aseev V.V. (1978), On a problem of rigidity of shells with curvilinear axis, *Prikladnaya mekhanika*, Vol. 14, № 5, 36-44.
- [57] Vainberg D.V., Gulyaev V.I. (1975), Stability of mechanical and physical fields in shells of complex form, *Uspehi mekhaniki deformir. sred*, Moscow: "Nauka", 96-104.
- [58] Panasyuk L.S. (1981), On choice of bending of helicoidal surfaces, *Prikl. geometriya i ing. grafika*, Kiev, Vol. 31, 96-98.
- [59] Whiston G.S. (1982), Use of screw translational symmetry for the vibration analysis of structures, *Inter. J. Num. Meth. In Eng.*, vol. 18, No 3, 435-444.
- [60] Aronson A.A., Zhubritskaya M.A., Sokolov V.V. (1991), A spiral chamber of turbines of Bureiskaya hydroelectric station, *Raschet predel. sostoyania beton i zhelesobet. konstruktziy energ. soor.*, PREDSO-90: Vses. nauchno-techn. soveshanie, Ust-Narva, May 22-24 1990, SPb, 123-129.
- [61] Dzhashiashvili T.G., Karagashev D.A. (1992), A method of calculation of natural frequencies of metal spiral chamber, *Issled. rats. i ekonom. konstruk. gidro i teploenerg. soor. dlya gornih usloviy*: GruzhnII energ. i gidrotehn. soor., 135-145.
- [62] Plotnikov F.A. (1964), An analysis of reinforced concrete spiral chambers of water turbines of circular cross-section, *Primenenie zhelesobetona v mashinostroenii*: Sb. Statey, Moscow: Mashinostroenie, 440-457.

Extraction of Conditional Functional Dependencies using Datamining Techniques

Jetty Rajesh*, V.Kumar**

*M.Tech Student, Avanathi Institute of Engineering and Technology, Narsipatnam, India

**H.O.D, Dept of CSE, Avanathi Institute of Engineering and Technology, Narsipatnam, India

ABSTRACT: Functional Dependencies (FD's) are recently extended as Conditional Functional Dependencies (CFD's) for cleaning Relational Data by supporting patterns of semantically related constants. Finding worth CFD's is an expensive process which needs more manual work. Using relations we are able to effectively identify data cleaning rules. Even the mining patter provided with new techniques, we in this paper provided three techniques for CFD discovery. First is a technique for closed mining item sets which is used for discovering constant CFD's ie's CFD's only with Patterns which Is essential to data cleaning and integration, which is referred to as CFDMiner. The other two Algorithms are CTANE and TANE which are developed for discovering General CFD's. And another Alogirhtmis FastCFD which is based on Depth-First Search Approach which will reduce the search space by identifying closed itemset mining. With our analysis, CFDMiner efficiently discovers constant CFDs. For general CFDs, CTANE works well when a given relation is large, but it does not scale well with the arity of the relation. FastCFD is far more efficient than CTANE when the arity of the relation is large; better still, leveraging optimization based on closed-itemset mining, FastCFD also scales well with the size of the relation. Thus we are providing the user to select set of cleaning rule discovery tools

Keywords: Privacy, Privelets, Data Publishing, Range count Queries.

I. INTRODUCTION

This paper investigates the discovery of conditional functional dependencies (CFDs). CFDs are a recent extension of functional dependencies (FDs) by supporting patterns of semantically related constants, and can be used as rules for cleaning relational data. However, finding CFDs is an expensive process that involves intensive manual effort. To effectively identify data cleaning rules, we develop techniques for discovering CFDs from sample relations. We provide three methods for CFD discovery. The first, referred to as CFDMiner, is based on techniques for mining closed itemsets, and is used to discover constant CFDs, namely, CFDs with constant patterns only. The other two algorithms are developed for discovering general CFDs. The first algorithm, referred to as CTANE, is a levelwise algorithm that extends TANE, a well-known algorithm for mining FDs. The other, referred to as FastCFD, is based on the depthfirst approach used in FastFD, a method for discovering FDs. It leverages closed-itemset mining to reduce search space. Our experimental results demonstrate the following.

(a) CFDMiner can be multiple orders of magnitude faster than CTANE and FastCFD for constant CFD discovery. (b) CTANE works well when a given sample relation is large, but it does not scale well with the arity of the relation. (c) FastCFD is far more efficient than CTANE when the arity of the relation is large.

As remarked earlier, constant CFDs are particularly important for object identification, and thus deserve a separate treatment. One wants efficient methods to discover constant CFDs alone, without paying the price of discovering all CFDs. Indeed, as will be seen later, constant CFD discovery is often several orders of magnitude faster than general CFD discovery. Levelwise algorithms may not perform well on sample relations of large arity, given their inherent exponential complexity. More effective methods have to be in place to deal with datasets with a large arity. A host of techniques have been developed for (non-redundant) association rule mining, and it is only natural to

capitalize on these for CFD discovery. As we shall see, these techniques can not only be readily used in constant CFD discovery, but also significantly speed up general CFD discovery. To our knowledge, no previous work has considered these issues for CFD discovery.

II. BACKGROUND & RELATED WORK

As remarked earlier, constant CFDs are particularly important for object identification, and thus deserve a separate treatment. One wants efficient methods to discover constant CFDs alone, without paying the price of discovering all CFDs. Indeed, as will be seen later, constant CFD discovery is often several orders of magnitude faster than general CFD discovery.

Levelwise algorithms may not perform well on sample relations of large arity, given their inherent exponential complexity. More effective methods have to be in place to deal with datasets with a large arity. A host of techniques have been developed for (non-redundant) association rule mining, and it is only natural to capitalize on these for CFD discovery. As we shall see, these techniques can not only be readily used in constant CFD discovery, but also significantly speed up general CFD discovery. To our knowledge, no previous work has considered these issues for CFD discovery.

In light of these considerations we provide three algorithms for CFD discovery: one for discovering constant CFDs, and the other two for general CFDs.

(Module: 1) we propose a notion of minimal CFDs based on both the minimality of attributes and the minimality of patterns. Intuitively, minimal CFDs contain neither redundant attributes nor redundant patterns. Furthermore, we consider frequent CFDs that hold on a sample dataset r , namely, CFDs in which the pattern tuples have a support in r above a certain threshold. Frequent CFDs allow us to accommodate unreliable data with errors and noise. Our algorithms find minimal and frequent CFDs to help users identify quality cleaning rules from a possibly large set of CFDs that hold on the samples.

(Module: 2) Our first algorithm, referred to as CFDMiner, is for constant CFD discovery. We explore the connection between minimal constant CFDs and closed and free patterns. Based on this, CFDMiner finds constant CFDs by leveraging a latest mining technique, which mines closed itemsets and free itemsets in parallel following a depth-first search scheme.

(Module: 3) Our second algorithm, referred to as CTANE, extends TANE to discover general CFDs. It is based on an attribute-set/pattern tuple lattice, and mines CFDs at level $k + 1$ of the lattice (i.e., when each set at the level consists of $k+1$ attributes) with pruning based on those at level k . CTANE discovers minimal CFDs only.

(Module: 4) Our third algorithm, referred to as FastCFD, discovers general CFDs by employing a depth-first search strategy instead of the levelwise approach. It is a nontrivial extension of FastFD mentioned above, by mining pattern tuples. A novel pruning technique is introduced by FastCFD, by leveraging constant CFDs found by CFDMiner. As opposed to CTANE, FastCFD does not take exponential time in the arity of sample data when a canonical cover of CFDs is not exponentially large.

(Module: 5) Our fifth and final contribution is an experimental study of the effectiveness and efficiency of our algorithms, based on real-life data (Wisconsin breast cancer and chess datasets from UCI) and synthetic datasets generated from data scraped from the Web. We evaluate the scalability of these methods by varying the sample size, the arity of relation schema, the active domains of attributes, and the support threshold for frequent CFDs. We find that constant CFD discovery (using CFDMiner) is often 3 orders of magnitude faster than general CFD discovery (using CTANE or FastCFD). We also find that FastCFD scales well with the arity: it is up to 3 orders of magnitude faster than CTANE when the arity is between 10 and 15, and it performs well when the arity is greater than 30; in contrast, CTANE cannot run to completion when the arity is above 17. On the other hand, CTANE is more sensitive to support threshold and outperforms FastCFD when the threshold is large and the arity is of a moderate size. We also find that our pruning techniques via itemset mining are effective: it improves the performance of FastCFD by 5-10 Folds and makes FastCFD scale well with the sample size. These results provide a guideline for when to use CFDMiner, CTANE or FastCFD in different applications.

These algorithms provide a set of promising tools to help reduce manual effort in the design of data-quality rules, for users to choose for different applications. They help make CFD-based cleaning a practical data quality tool.

III. SYSTEM DESIGN & ANALYSIS

This Component design diagram helps to model the physical aspects of an object oriented software system i.e., for the proposed framework it illustrates the architecture of the dependencies between service provider and consumer.

The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted

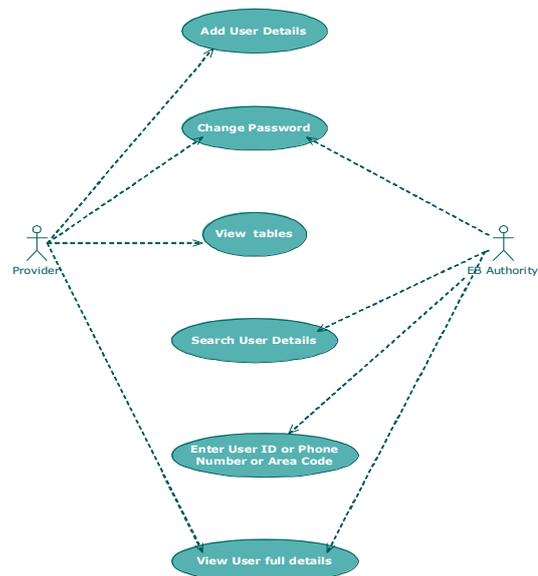


Fig.1: Inter-operational Use case diagram for the framework

A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner

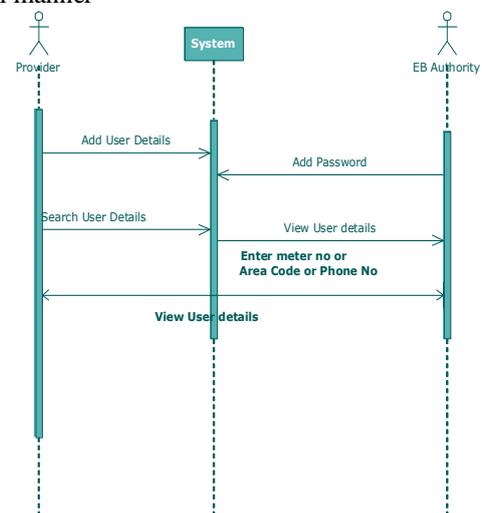


Fig.2: Inter-operational Sequence Diagram for the Framework

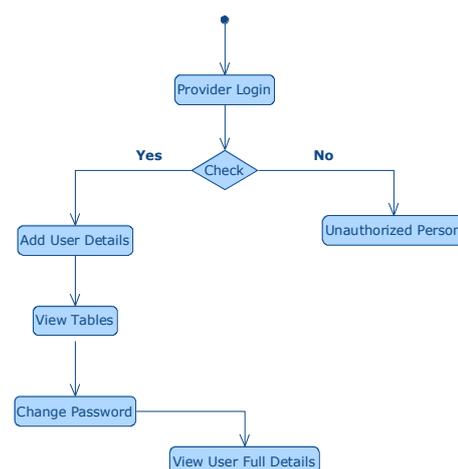


Fig.3: Inter-operational sequence diagram for Framework

RESULTS

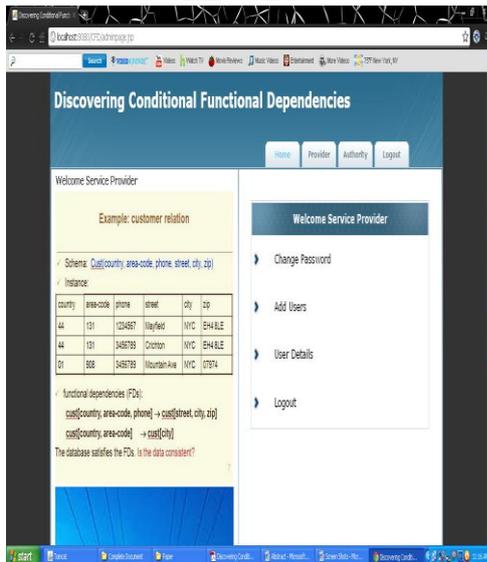


Fig.4: To add the user Details

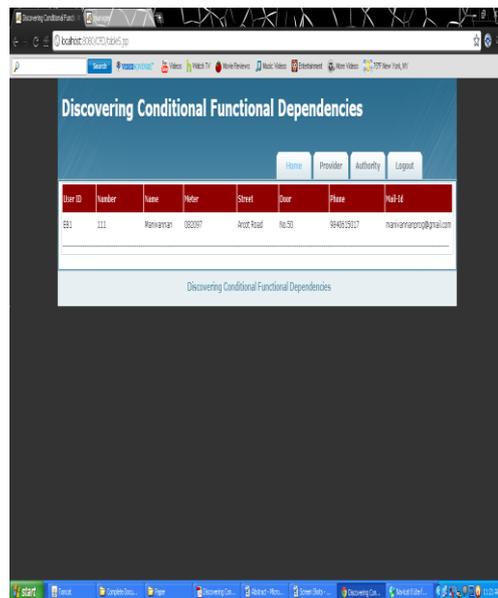


Fig.7: User complete information

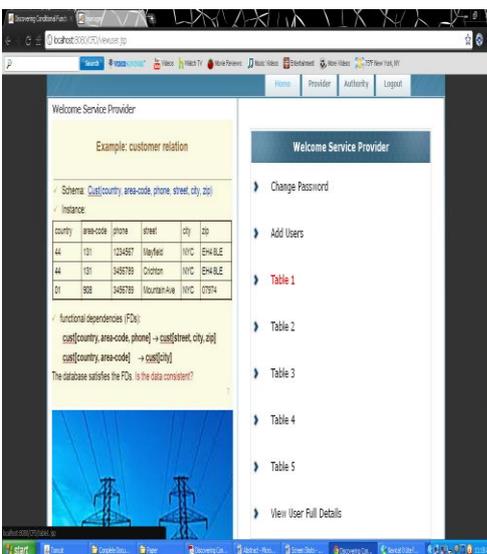


Fig.5: To give the Table information

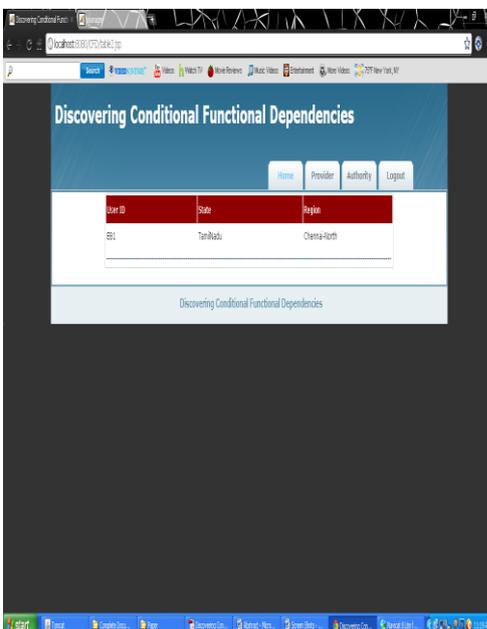


Fig.6: User output Screen

CONCLUSION

We have developed and implemented three algorithms for discovering minimal CFDs: CFDMiner for mining minimal constant CFDs, a class of CFDs important for both data cleaning and data integration; CTANE for discovering general minimal CFDs based on the levelwise approach; and FastCFD for discovering general minimal CFDs based on a depth-first search strategy, and a novel optimization technique via closed-itemset mining. As suggested by our experimental results, these provide a set of tools for users to choose for different applications. When only constant CFDs are needed, one can simply use CFDMiner without paying the price of mining general CFDs. When the arity of a sample dataset is large, one should opt for FastCFD. When k-frequent CFDs are needed for a large k, one could use CTANE.

REFERENCE

- [1] L. Bravo, W. Fan, F. Geerts, and S. Ma, "Increasing the expressivity of conditional functional dependencies without extra complexity," in ICDE, 2008.
- [2] G. Cormode, L. Golab, F. Korn, A. McGregor, D. Srivastava, and X. Zhang, "Estimating the confidence of conditional functional dependencies," in SIGMOD, 2009.
- [3] L. Bravo, W. Fan, and S. Ma, "Extending dependencies with conditions," in VLDB, 2007.
- [4] B. Goethals, W. L. Page, and H. Mannila, "Mining association rules of simple conjunctive queries," in SDM, 2008.
- [5] S. Lopes, J.-M. Petit, and L. Lakhali, "Efficient discovery of functional dependencies and armstrong relations," in EDBT, 2000.
- [6] T. Calders, R. T. Ng, and J. Wijsen, "Searching for dependencies at multiple abstraction levels," TODS, vol. 27, no. 3, pp. 229-260, 2003.
- [7] R. S. King and J. J. Legendre, "Discovery of functional and approximate functional dependencies in relational databases," JAMDS, vol. 7, no. 1, pp. 49-59, 2003.
- [8] I. F. Ilyas, V. Markl, P. J. Haas, P. Brown, and A. Aboulnaga, "Cords: Automatic discovery of correlations and soft functional dependencies," in SIGMOD, 2004.
- [9] H. Mannila and H. Toivonen, "Levelwise search and borders of theories in knowledge discovery," Data Min. Knowl. Discov., vol. 1, no. 3, pp. 259-289, 1997.
- [10] Gartner, "Forecast: Data quality tools, worldwide, 2006-2011," 2007.
- [11] S. Abiteboul, R. Hull, and V. Vianu, Foundations of Databases. Addison-Wesley, 1995.
- [12] L. Golab, H. Karloff, F. Korn, D. Srivastava, and B. Yu, "On generating near-optimal tableaux for conditional functional dependencies," in VLDB, 2008.

Implementation of Zero Knowledge Protocol for Anomaly based Intrusion Detection and Prevention System in Wireless Sensor Network at Local and Remote Level

Manish P. Gangawane¹, Santhosh Easo², Priya Balkrishna Malve³

M.E Student, Department of I.T, M.I.T.M, Indore, R.G.T.U. University, Bhopal

Associate Proff. Comp-Sci dept, M.I.T.M, Indore R.G.T.U. Bhopal

Lecturer, Computer Engg. Dept. S.V.I.T. Chincholi, University Of Pune

ABSTRACT: Wireless Sensor networks (WSN) is excellent technology which provide great potential for situations like battlefields and commercial applications such as building, traffic survey, monitoring environments smart homes and many more scenarios. Security is the most important challenge in wireless sensor networks. Sensor networks dose not have any user control for each individual node and wireless environment. Basically some special security threats and attacks of WSNs get addressed in our paper. Distributed sensor cloning attack will get identified using this model. We implement zero knowledge protocol (ZKP) for the verification of sender sensor nodes. With attachment of unique fingerprint to each node we address the clone attack. In the wireless sensor network non transmission of crucial cryptographic information is addressed by our model using ZKP. So it is helpful for preventing man-in-the middle attack and replay attack. Detailed information about different scenarios and also analysis of performance and cryptographic strength are content of this paper.

Keywords: clone attack, man in middle attack, replay attack, anomaly based intrusion detection, zero knowledge protocol, WSN.

I. INTRODUCTION

Because of the advanced technologies available now a days, so it possible to develop the sensor node in wireless sensor networks. Basically these kinds of nodes are compact and they are attaching with a variety of sensors and mostly wireless. Minimal manual intervention and monitoring is done after the deployment of these sensor nodes. But, there may be issues of security concern as we deploy the nodes in the hostile environment and where there is no manual controlling of nodes. Normally clone nodes in the network are one of the most important type of physical attack. It is easy for adversary to identify the authorized nodes, cryptographic information copy to make clones and these clones are deployed back to the network by using commodity hardware and operating system. The hardly appropriation of general purpose security protocol is due to these constraints. The main aim of the paper is to implementation of a security model for wireless sensor networks and to classify various attacks of it. Man In Middle attack, Clone attack and Replay attack of WSN's are easily get identified by this method and also verification of authorized sender sensor nodes in wireless sensor network for this we uses zero knowledge protocol. Using intrusion detection and prevention system administrator are able to block or kill any suspicious activity which found in system, The anomalous activity

scanning and killing is done by administrator locally or even remotely by using the GPRS facility of his mobile device.

II. SECURITY GOALS FOR WIRELESS SENSOR NETWORKS

Primary and secondary are the main types of security goals are there in Wireless Sensor Network. The primary goals are known as standard security goals such as Confidentiality, Integrity, and Authentication. The secondary goals are Data Freshness, Time Synchronization and Secure Localization. These goals are explained as follows. Primary goals are as:

A. Data Confidentiality

In sensor network the ability to conceal messages from a passive attacker is confidentiality. Due to this message communication through sensor network remains confidential. A sensor node should not shows its data to the neighbors.

B. Data Authentication

The reliability of the message through identification of its origin done by authentication. Alteration of packets are basically involves in attacks of WSN Identification of senders and receivers are verified by data authentication

C. Data Integrity

Data reliability is insured by Data integrity in sensor networks. It also have an ability that confirm message has not been tampered with, altered or changed. Secondary goals are

D. Secure Localization

A sensor network designed to ensure faults. It accurate information related with location for identification of location fault.

III. PREPARATORY

In this section, we introduce the basics of s-disjunct code, which incorporates social characteristics and used to generate fingerprint for each sensor node [1]. These fingerprints are subsequently used to detect clone attack. Let \mathbf{X} be a $m \times n$ binary matrix. Matrix \mathbf{X} is considered as constant column weight ω and a constant row weight λ . Then, $mX_{i,j} = \omega$ and $nX_{i,j} = \lambda$ Where $1 \leq i \leq m$, $1 \leq j \leq n$. The binary matrix \mathbf{X} can be used to define a binary codeword, with each column $\mathbf{X}_j = (X_{1,j}, X_{2,j}, \dots, X_{m,j})^T$

Definition 1 Given two binary codeword's $\mathbf{y} = (y_1,$

$y_2, \dots, y_m)^T$ and $\mathbf{z} = (z_1, z_2, \dots, z_m)^T$. [12]

Definition 2 An $m \times n$ binary matrix X defines a superimposed code of length m , size n , strength s ($1 < s < m$), and list size L ($1 \leq L \leq m - s$), if the Boolean sum of any s -subset of columns of X can cover no more than L columns of X which are not in the s -subset. This code is also called as (s, L, m) -code of size n . [7]

Definition 3 A binary matrix defines an s -disjunct code if and only if the Boolean sum of any s -subset of columns of X does not cover any other column of X that are not in the s -subset. As per the s -disjunct characteristic of superimposed s -disjunct codes, important property follows, can be employed to compute fingerprints to detect clone attacks. [13] **Property 1** Given a superimposed s -disjunct code X , for any s -subset of columns of X , there exists at least one row in X that intersects all the s columns with a value 0. Generation of a good superimposed s -disjunct code has been extensively studied in literature ([6,7,9,14]).

IV. IMPORTANT ATTACKS IN WSN

Number of security attacks there in wireless sensor networks. But our proposed model can detect certain attacks as follows:

A. Man in the Middle Attack

In man-in-the-middle attack (MITM) an attacker sits between sender and receiver and sniffs any information being sends between them.

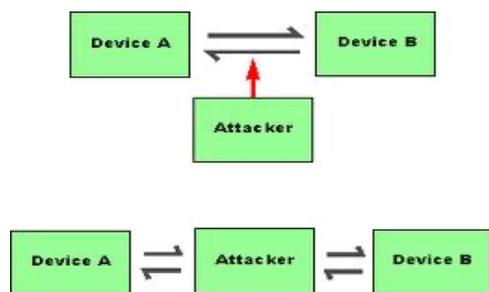


Fig 1 Main in Middle attack

In this third party makes independent connections with the victims and messages send between them. Due to this the sender and receiver thinks that they are talking directly with one other private connection.

B. Clone Attack

Sensors are susceptible to physical capture attack is the one of the most susceptible issue in wireless sensor network. After the compromisation of sensor the adversary can easily launch clone attack by replicating the compromised node. After this it distributes the clones to entire network and starting the variety of insider attacks. In detection of cloning attack continuous physical monitoring of nodes is impossible.

C. Replay Attack

The already sent packets are repeats the malicious node is the replay attack. Due to this it results in nodes energy

exhaustion thus network get collapse.

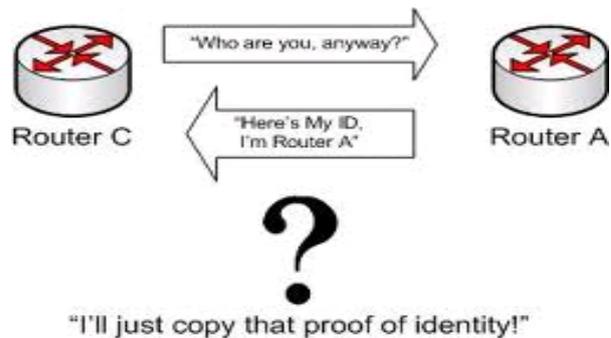


Fig 2: Replay Attack

D. Anomaly Based Intrusion Detection

This system we focus on to acquiring volatile data which leave no trails once the system is power off. The volatile data can be in the form of RAM Contents, temporary data used by the OS, data in registers, buffers, unlinked file and unsaved files; and these volatile data may contains information about all running processes, active and recent network connections, open ports and sockets, processes running in background, open files and applications, loaded DLLs, OS kernel module, and active users. These volatile data can have enough information about the anomalous activities on running system.

V. ZERO KNOWLEDGE PROTOCOL

Authentication systems motivates all the research of zero knowledge proofs in which prover wants to prove its identity to a verifier through some secret information (such as a password) but never wants that the second party to get anything about this secret known as "zero-knowledge proof. Identification, key exchange and other cryptographic operations is mainly allowed by Zero Knowledge Protocol. Implementation is done without showing any important information during the conversation. For resource constrained devices ZKP is very useful and attractive. ZKP is an interactive proof system which involve node P and node V. P plays prover role where as V is verifier. In a series of communications prover conveys the verifier of some secrets through series of communication. In each and every communication a challenge, or question, are comes from the verifier and basically prover response. Normally less bandwidth, less, small computational power, and less memory is needed by ZKP based protocols.

A. Mechanism Of ZKP

In WSN using ZKP one party assures another that a statement is true instead of showing anything other than the veracity of the statement. The prover and the verifier uses some numeric value, which acts as secret number for prover. Basically computational intensive mathematical problem are normally offered by prover p, and many possible solutions to this problem are normally requested from verifier side. If critical information relating to the solution is knows by p then replay with any one requested available solutions to it. If the P not knows anything related with critical information, then he is enable to provide the needed information to the V.

VI. CRYPTOGRAPHIC STRENGTH

The cryptographic strength of ZKP is based on hard to solve problem. We uses problem of factoring large numbers which are product of two or more large prime numbers. As the value of public key get changed with every communication so it is not easy for attacker to identify it. The prover also generates a random number and the fingerprints also changes randomly. Thus as public key changes challenge question from verifier and a new random number from the prover, becomes extremely difficult for the attacker to break the security.

VII. PROPOUND MODEL

A. Assumptions

Base station, cluster head and member nodes are three main nodes in this model. Mostly random nodes are considered as cluster heads. Each and every cluster head had information about its member nodes and vice versa. The information about all sensor nodes which includes cluster heads also is stored in base station. Base station, maintains all the topological information about cluster heads and their respective members by communication among member nodes is not possible.

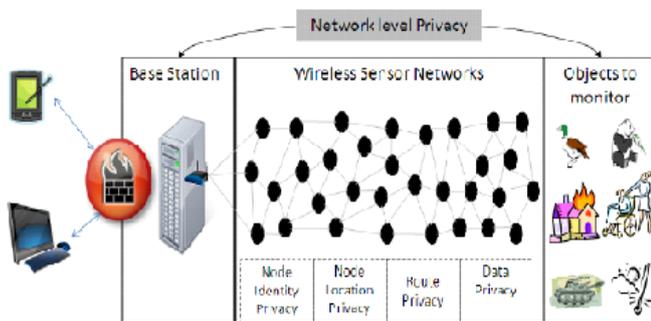


Fig. 3. Communications in the proposed model

B. Pre-deployment phase

For deploying the nodes in the network, we generate a unique fingerprint for each sensor node. It addressed by combining relative nodes information through a superimposed s-disjunct code and this is preloaded in each node. Due to this each node seems unique from other one. Basically this fingerprints remains secret throughout the process.

C. Post-deployment Phase

A public key N generation by the base station is done after the deployment. Basically this key is used by any two nodes at a given time while communicating. Here base station is third party whereas sender node is prover and receiving node verifier. Each node is assigned a fingerprint which is used as a private key (secret key). Prover and receiver shares the public key. Now from base station secret key of the prover from the base station is requested by verifier. The base station will generate a secret code $v = s2modN$ (where s is finger print of the prover and N is the public key). The value of v is given to the verifier on its request [13]. Fingerprint is never shown or transmitted in the network directly during this entire communication process. By using ZKP for k times per communications verifier will continues the authentication process which includes number

of verification rounds. Failure of prover for authentication of itself in any one of the k rounds, then it becomes a compromised node. For more effectiveness of protocol it must be passed through large number of rounds. The number s remains private within the domain of the prover. Thus makes it computationally infeasible to derive s from v given $v = s2modN$.

VIII. ANOMALY BASED INTRUSION DETECTION SYSTEM

The analysis is the heart of the anomaly intrusion detection system. In this system we investigate user patterns, such as profiling the programs executed daily or the privileged processes executed with access to resources that are inaccessible to ordinary user. For this we collect the volatile data from the system. To collect this data we use system log file which gives us the number of processes which are running on the system, which are provided to the user, and for privileged of system. We trained our system by using conditional random fields, which reduces the false alarm rate. Then the system is deployed in real working environment. If the anomalous activity occurs then we alerts the administrator by sending SMS that the anomalous activity is running.

IX. IMPLEMENTATION

In this project we are going to create wireless sensor network which belongs one server and multiple client then identifying various attacks in WSN by ZKP. Client can be registered to network for this facility we need to do java RMI programming. Both the client and server side will communicate by using the zkp protocol. Attacker will try to perform attack on the network all the log will be captured by zkp prevention system admin can take action depend on the log obtained on server. If we found any intrusion in the network then we can prove that our communication will not affected by intruder. Central database will be developed in Microsoft access which will communicate with all the node by using DSN. All the client will have login facility along with new registration. The communication will be displayed on the system by using swing frame. The logs are dynamic as the system changes it will show different record relevant to the current situation parameter. To develop this system we need networking socket programming and database programming. Using ZKP it is easy for us to indentify the attacks in WSN.

A. Prevention of Anomalous Activity

Once the anomalous activity occurs, we can prevent it. The admin may log on to the system locally or remotely. If the administrator is at local level then he/she can view the running activities, or he/she can stop the anomalous activity and if the administrator is at remote level then he/she can log on to the system using Internet or GPRS using cell phone. After that the user can stop anomalous activity, or start new activity. But if the controlling of the anomalous activity is not possible then administrators may shutdown or reboot that system.

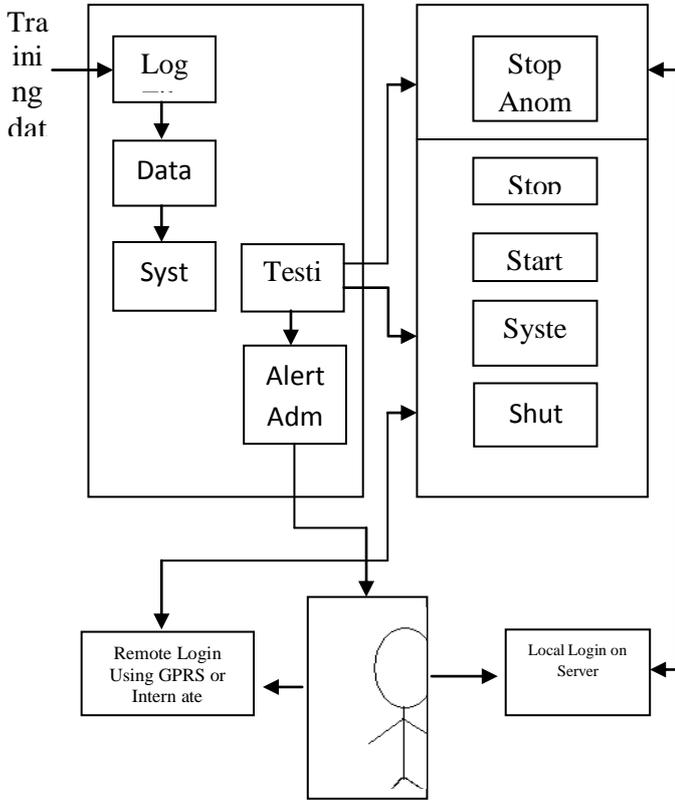


Fig 4 Prevention of Anomalous Activity

B. Generation of unique fingerprint for each node

The base station is assumed to be aware of the topology of the network and all neighborhood information. Before deployment, the base station computes the finger print for each node in the network. For every node u , base station finds its neighborhood information. In our approach, the neighbourhood $Ngh(u)$ should satisfy $ng < s$, where ng is the number of sensor nodes in $Ngh(u)$, s is the strength of the superimposed code X . Finger print for sensor node u is computed by considering the code words of all node v which are in the $Ngh(u)$. Given a sensor node u , base station computes u 's fingerprint as follows. Let $Xu = Xu1, Xu2, \dots, Xung$ denotes the codeword set of the nodes in $Ngh(u)$, where $Xu i$ denotes the codeword of u 's i -th closest neighbor[13]. Out of all Xu , the boolean sum of s -closest neighbors of node u ($Xu s$), is computed first. According to the property of the superimposed s -disjunct code, the resulting vector should contain at least one element with a zero value. These zero elements making relationship among neighbors s , that actually represent the social characteristic of sensor node u . Motivated by this observation, we use binary representation of the position of a zero element in the boolean sum of $Xu s$ as the social fingerprint of u . Intuitively, the social fingerprint should be stronger if more information from $Ngh(u)$ is brought in during the fingerprint computation [1].

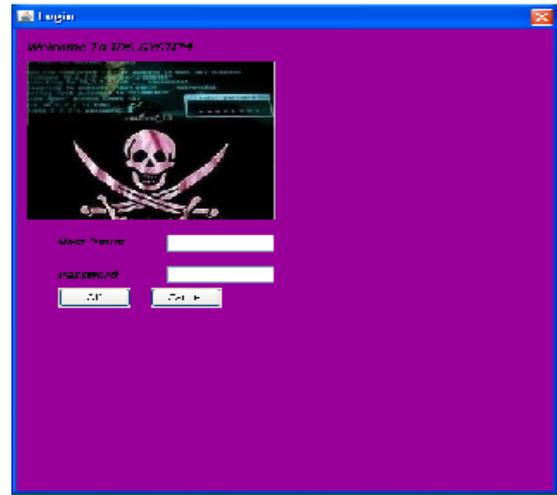


Fig: 4 Intrusion Detection System

Base station repeats this procedure mentioned in figure 2 to compute the fingerprint for each node u in the network [1]. The method starts with a s -subset of $X(u)$ that contains the code words of the s closest neighbors of sensor node u , and expands the subset until any further increment will not have a zero element in the boolean sum. For the subset resulting from the last increment, boolean sum is computed and position of one of the zero elements in the resulting sum get select. The binary equivalent of this position value is denoted as the finger print of node u . By taking u 's Id as seed for the pseudo random function, base station is able to compute unique positions for zero element [1].

X. SECURITY ANALYSIS OF PROPOSED MODEL

A. Cloning Attack

Case 1: Any other existing id with same fingerprint get used by cloned node: As an node get compromised its clones are inserted to network which always tries to make a part of communication. Only after the verification of clone nodes they are able to communicate with other nodes Fig 5 shows how node '6' of cluster '2' is get cloned and placed in cluster '1' with a new id '2'. Cloned node uses the fingerprint 's' of node '6', it fails to authenticate itself during communication through ZKP.

Case 2: When same id and same fingerprint used by cloned node:

If it uses the same id '6', the cluster head of cluster 1 will reject any communication as node '6' as it is not a member of cluster '1'. The base station which will detect immediately at the initiation of the communication request. This scenario is depicted in

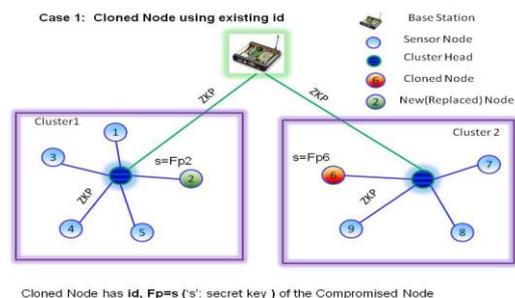


Fig.5. Case1 of security analysis

Case 3: When already present id with a different finger print get used by clone nodes: The cloned node with some existing Id get detected every time by the neighboring nodes (cluster heads) as the secret finger print of the cloned node will not match with the finger print possessed by the neighbors.

Case 4: When a cloned node behaves as a cluster head The cluster heads communicate with base station. The base station becomes the verifier and poses the challenge question to the cloned cluster head and detects the cloning attack through ZKP.[13]

B. Man in Middle Attack

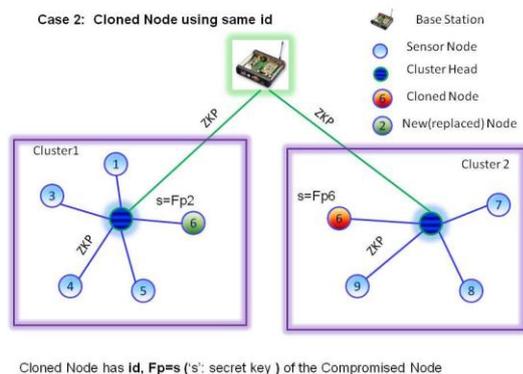


Fig6: Case 2 of security analysis

In our model, the finger print of a node never gets transmitted and thus intruder not has chance to identify them. Even if the attacker tries to generate a finger print in some brute force method, it will not be able to escape the check as every time a new public key N and a new random challenge question will be used.

C. Replay Attack

In this attack, an intruder tries to replay the earlier communication and authenticate itself to the verifier. But, with our model verifier will be sends different values each and every time in communication, replaying earlier communication.

XI. CONCLUSIONS

This paper proposed a new security model which addresses three important types of active attacks MITM attack, Clone attack and Replay attack. By using Zero knowledge protocol we implement this model. The proposed model uses finger print for each and every communication between the node. Thus it is easy for the administrator to identify these attacks using ZKP. Different types of attack there related information, different cryptographic strength and performance of the proposed model get analyzed in this system.

REFERENCES

[1] Kai Xing Fang, Liu Xiuzhen, Cheng David, H. C. Du, Real-Time Detection of Clone Attacks in Wireless Sensor Networks Proceedings of the 28th International Conference on Distributed Computing Systems, 2008, Pages 3-10.

[2] Joseph Binder, Hans Peter Bischof, Zero Knowledge Proofs of Identity for Ad Hoc Wireless Networks An In-Depth Study, Technical Report, 2003. <http://www.cs.rit.edu/jsb7384/zkp-survey.pdf>

[3] Klempous R.; Nikodem J.; Radosz, L.; Raus, N. *Byzantine Algorithms in Wireless Sensors Network*, Wroclaw Univ. of Technol., Wroclaw; Information and Automation, 2006. ICIA 2006. International Conference on, 15-17 Dec. 2006, pages:319-324

[4] I. Krontiris, Z. Benenson, T. Giannetos, F. C. Freiling, and T. Dimitriou, Cooperative Intrusion Detection in Wireless Sensor Networks, in Proc. EWSN'09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 263-278.

[5] A. J. Macula. A simple construction of d-disjunct matrices with certain constant weights Discrete Math., 162(13):311-312, 1996

[6] K. Xing, X. Cheng, L. Ma, and Q. Liang., Superimposed Code Based Channel Assignment in Multi-radio Multi-channel Wireless Mesh Networks. In MobiCom'07, pages 15-26, 2007.

[7] Md. Moniruzzaman, Md. Junaid, Arafeen, Saugata Bose, Overview of Wireless Sensor Networks: Detection of Cloned Node Using RM, LSN SET, Bloom filter and AICN Protocol and Comparing

[8] H. Choi, S. Zhu, and T. Laporta., Set: Detecting Node Clones in Sensor Networks. In SecureComm'07, 2007.

[9] Goldreich, O., Micali, S., and Wigderson, Proofs That Yield Nothing But Their Validity Or All Languages in NP Have Zero knowledge Proof Systems, Journal of the ACM, Vol. 38, No. 1, pp.691-729, 1991.

[10] Tuyls, Pim T. (Mol, BE), Murray, Bruce (Eastleigh GB), Efficient Implementation of Zero Knowledge Protocols, United States NXPB.V. (Eindhoven, NL) 7555646, June 2009, <http://www.freepatentsonline.com/7555646.html>.

[11] Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in Wireless Sensor Networks," IEEE Commun. Surveys Tutorials, vol. 8, pp. 2-23, year 2006. (IJCSIS) International Journal of Computer Science and Information Security, Vol. 4, No. 1 & 2, 2009.

[13] Siba K. Udgata, Alefiah Mubeen, Samrat L. Sabat Wireless sensor security model using zero knowledge protocol, 978-1-61284-233-2/11/\$26.00 ©2011 IEEE.

[14] A. G. Dyachkov and V. V. Rykov., Optimal superimposed Codes and designs for Renyis Search Model. Journal of Statistical

[15] A. Vanathi, B. Sowjanya Rani, Cloning Attack Authenticator In Wireless sensor network, IJCST Vol. 3, Issue 1, Sp 1. 5, Jan-Mar 2012 ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

[16] Sandip A. Shivarkar, Mininath Raosaheb Bendre, Hybrid Approach For Intrusion Detection Using Conditional Random Fields. International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 1, Issue 3

[17] Annarita Giani, Security Essentials - Version 1.2e, Identification With Zero knowledge protocol, © SANS Institute 2001.

[18] Hannu A Aronsson, Zero Knowledge Protocols and Small systems.

[19] Shivarkar Sandip, Mujumdar Ajit, Dange Bapusaheb j, An Improved Approach For Anomaly Based Intrusion Detection And Prevention.

[20] Vishal Prabhat, Tushar Manikrao, Nitesh tayade, Sushila Aaghav, Zero knowledge protocol to design security model for threats in WSN, ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 2, Mar-Apr 2012, pp.1533-1537.

A Study on the Ductility of Bolted Beam- Column Connections

Shemy S. Babu¹, S. Sreekumar²

Department of Civil Engineering College of Engineering, Trivandrum, Kerala

ABSTRACT: The behaviour of beam-column joint plays an important role in the response of a steel moment resisting framed structure especially under seismic excitation. Semi rigid connections like bolted joints, allow better energy dissipation and provide good response under lateral loads. In this paper the ductility and energy dissipation characteristics of semi rigid bolted connections using double web angle with top and seat angles is discussed based on experimental investigation. Non-linear analysis was also performed using finite element method to compare the results. The study revealed that ductile behaviour of beam-column connection is improved by increasing the number of bolts on the top and seat angles and the thickness of top and seat angles. The analytical results are in agreement with the experimental results.

Keywords: semirigid connections; beam column joints

I. INTRODUCTION

Structural steel framework with welded joints was considered as one of the best moment resisting framed structural system till the 1994 Northridge earthquake. In the earthquake, many structural steel frames failed due to the occurrence of brittle fractures at the connections. This caused the loss of reliability on steel welded moment resisting frames. Since then, as an alternative, bolted connections, often called semi-rigid connections, are considered for design of steel moment resisting frames and retrofitting works.

The behaviour of beam-column joint plays an important role in the response of a steel moment resisting framed structure. It strongly influences the seismic behaviour and energy dissipation capacities of the moment resisting frames. The use of bolted joints allows more energy dissipation and provides a good response under lateral loads. In structural framework with semi-rigid joints, the characteristics of connections play a significant role in the energy-dissipation mechanisms. As rigid connections are more expensive and difficult to assemble and because flexible connections do not have the necessary resistance and stiffness to resist the lateral loads, use of semi-rigid connections using bolted angles is justified. Semi-rigid connections with angles connecting the web and flanges of the beam to the column flange are cheaper and simpler to assemble when compared to other conservative joint.

Various studies have revealed that semi-rigid steel frames have the potential to be used in buildings located in high seismic prone regions. It was found that bolted-angle connections have stable cyclic response and energy dissipation capacity depending on the size of the angles and bolts.

This paper is concerned with the study of structural response of a typical semi-rigid steel beam-column connection. In particular, the behaviour of double web angle bolted beam-column connection is experimentally investigated under monotonic loading using top and seat

angle sections. Static half cycle load tests were performed to compare the energy absorption capacity of various connections. The effect of various joint parameters such as diameter of bolts, number of bolts on top and seat angle, thickness of top and seat angles etc. on energy dissipation capacity of the joint was studied experimentally. Analytical study on the behaviour of double web angle connections with and without top and seat steel angle sections simulating the experimental investigation was performed with the help of finite element software.

II. RELATED STUDIES

In recent years, a number of research works were carried out to study the behaviour of semi-rigid connections. This section deals with the review of the works related to the semi-rigid connections. Cheol and Young [1] conducted experimental investigation on the cyclic load behaviour of steel frames with bolted joints in comparison with welded joints for an exterior beam column joint. Three different types of connections viz. semi-rigid connection using high strength bolts, simple bolted connection and welded connection were considered in the study for varying effect of initial stiffness, maximum flexural resistance, failure modes, ductility and energy dissipation.

Chen et al. [2] discussed a new method of trimming the beam flanges in order to enhance the ductility of beam-column connection keeping the ultimate strengths unaltered and increasing the plastic rotation capacity and energy dissipation capacity.

Secer and Ozturk [3] studied the dynamic characteristics of semi-rigid frames and the influence of connection flexibility on the dynamic characteristics of frames analytically. Christopoulos et al. [4] presented characteristics of joints with post-tensioned energy dissipation connections. The beam-column connections were modelled as rotational springs attached at the ends of the members and the load carrying capacity, end member forces and deformation of frames with semi-rigid connections were determined using direct stiffness method. The influence of partial rigidity of beam-column connection on the behaviour and stability of frames was discussed (Frederick and Abuyasein, [5]). An expression for fixity factor, indicating the degree of rigidity of joint was derived in the study. Kartal et al. [6] conducted numerical studies for finding the response of frames with semi-rigid connections under different material and geometric idealization strategies.

III. EXPERIMENTAL INVESTIGATION

A schematic cantilever type test set up shown in Fig. 1 was adopted to study the semi rigidity behaviour of bolted beam-column joints. Ends of the column were supported and load was applied at the free end of beam member.

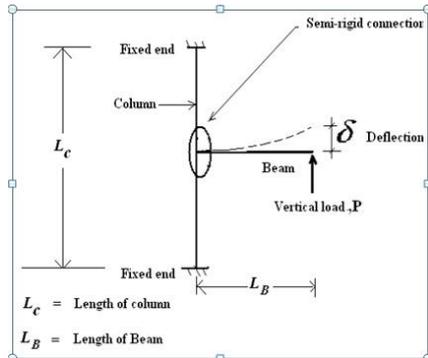


Fig. 1. Schematic diagram- Cantilever test set-up

The load was applied gradually and the deflections at three different points on the beam were measured. The cantilever test setup was arranged within the straining unit of 100t AVERY make universal testing machine. The column member was supported in the top and bottom compression plate of universal testing machine. Since the study was focused on the behaviour of connection, the use of convenient dimensions and boundary conditions adopted was justified. Considering the limitations in the space between the compression plates and connection facilities, suitable height and appropriate section was assumed for the column member. Standard steel section ISMB 200@25.4kg/m was used for beam and column members. The schematic diagram of the test setup is shown in Fig.2.

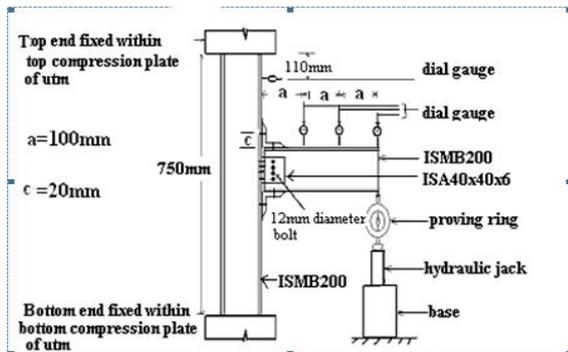


Fig.2. Schematic diagram-test set-up

The column was set free of initial stresses. The load was applied using a hydraulic jack and a proving ring of capacity 100kN was used for measurement. The dial gauges were fixed on the top of the beam member at the free end at a distance of 100mm, and 200mm from free end. A photograph of the experimental test setup is shown in Fig. 3. High strength structural bolts were used for connections.

The geometrical parameters varied in the study are the thickness of top and seat angle, number of bolts and the bolt diameter. The connection details of double web angle are given in Fig. 4. A gap of 6mm was provided between the column and beam to allow the rotation without damaging the column and beam members.

The ultimate condition was assumed as the closure of gap provided between the beam and column members or the failure of bolts whichever is earlier. The deflections at three points on the top of beam as shown in Fig. 2 were measured using dial gauges.

Table. I. Details of specimen tested

Specimen	Angle Sections	Bolt Diameter, d(mm)	Number of Bolts, n
S1	40X40X6	6	2
S2	40X40X6	10	2
S3	40X40X6	6	4
S4	35X35X3	6	4
S5	35X35X3	10	2
S6	35X35X3	6	2
S7	no angles at top and bottom double web angle connection		

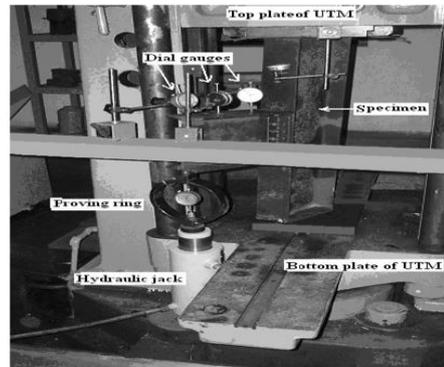


Fig. 3. Experimental setup

The connection details of the specimens tested are given in Table. I. The test specimen S1 had the connection with top and seat standard angle section ISA 40x40x6 using 2 numbers of 6mm diameter high strength bolt.

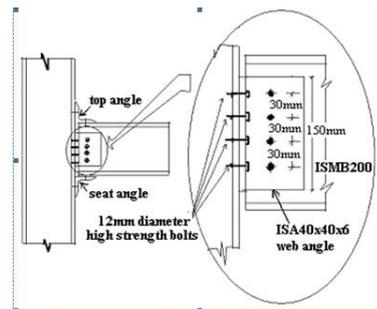


Fig. 4. Typical connection. (Inset shows the details of web angle connection).

IV. DUCTILITY

Ductility plays an important role in seismic design. The joint behaviour has a direct influence on the ductility. The ductility factor is one of the parameters characterising the behaviour of a structure or joint in plastic range. Considering the rotation of joints, rotational ductility factor is given by (1).

$$\mu_r = \frac{\phi_u}{\phi_y} \tag{1}$$

ϕ_u is the rotation on ultimate state and ϕ_y is the rotation on reaching first yield point of joints.

ENERGY DISSIPATION CAPACITY

The connection stiffness F of beam is related with moment (M) and rotation of beam at the joint (θ) by (2). The connection stiffness of column assuming the connection to be at the centre of column given by (3) can be

obtained from the analysis of a member fixed at both ends and subjected to a moment at the centre.

$$F = -\frac{M}{\theta} \quad (2)$$

$$\phi = -\frac{ML_C}{16EI_C} \quad (\text{anti-clockwise}) \quad (3)$$

ϕ is the rotation of the column at the joint, L_C is the length of column member, E is the Young's modulus of elasticity of the material, and I_C is the moment of inertia of column section.

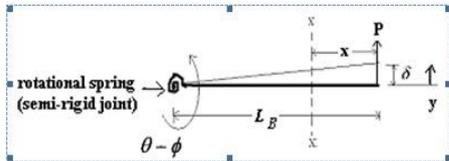


Fig. 5 Semi-Rigid Beam Member

The deflection of a member with one end free and other end connected semi-rigidly and subjected to point load at free end as shown in Fig. 5 is given by (4). I_B is the moment of inertia of the beam cross section. The clockwise rotation is assumed to be positive and downward deflection is considered negative.

$$\delta = \frac{PL_B^3}{3EI_B} + \frac{PL_B^2}{F} + \frac{PL_B^2 L_C}{16EI_C} \quad (4)$$

From (4) the beam connection stiffness, F , can be determined. Thus the beam rotation and column rotation can be determined by using (2) and (3). Using θ and ϕ , relative rotation at joint

($\theta_y = \theta - \phi$) can be computed. The moment-

relative rotation curve can be then plotted. The energy dissipated by the connection is quantified by computing the area under the moment-relative rotation curve.

V. FINITE ELEMENT ANALYSIS

In the present problem, moment-relative rotation behaviour of bolted beam-column connection is studied numerically. As the members associated with the connection undergo large deformations and rotations, geometric and material non-linearity are to be accounted in the problem. Since all the experiments are performed under static loading condition, a static non-linear structural analysis is required for simulating the real problem. Conventional 3D brick elements are inadequate to represent the realistic contact behaviour in the column-beam interface and bolt-plate interface under incremental loading conditions. Contact problems can be efficiently handled with ANSYS.

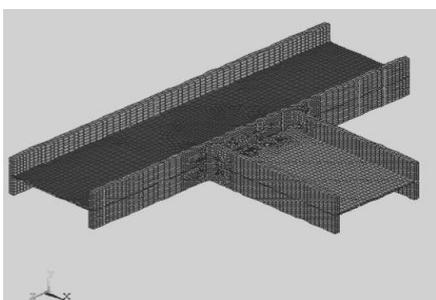


Fig. 6. Column-beam finite element mesh

Brick elements with 20 nodes, SOLID 95 available in the ANSYS element library, were used in the three dimensional modelling of beam, column angles and bolts. The contact surfaces including the areas anticipated to be in contact were defined and paired using contact elements, CONTA174 and TARGE170. For structural steel used, the yield stress was taken as 250 N/mm², ultimate tensile strength as 410 N/mm² and % elongation of 23 as per IS 2062:1999. For the high strength bolts used, yield stress of 640 N/mm², ultimate stress of 800 N/mm², and 12% elongation were adopted as per IS 1367 (Part 3):2002. The modulus of elasticity of steel was taken as 2x10⁵ N/mm². The column and beam members were modelled and mapped meshing consisting of brick elements was performed as shown in Fig. 6.

VI. RESULTS AND DISCUSSION

Experimental Results

The moment-rotation behaviour of specimens with different geometric parameters of double web angle connection with top and seat angle sections discussed is depicted in Fig. 7. In the figure, two regions, elastic region followed by yielding, can be noted for all repeated test cases. The curves of specimens S1 and S2 shown in Fig. 7 indicates that as the diameter of bolt (d) is increased, ultimate moment and ultimate rotation increased. There was an increase of 67% in the ultimate moment when the diameter was increased from 6mm to 10mm.

Table. II shows the effect of the ratio of diameter of bolt (d) to thickness of angle (t) on energy dissipation. When number of bolt was doubled an increase of 508% in energy dissipation was observed for top and seat angle connection with d/t ratio 1 and an increase of 62% for top and seat angle connection with d/t ratio 2. As the number of bolts increased (specimens S1 and S3) there is an increase in the ultimate rotation with slight increase in ultimate moment. Hence the ultimate moment and ultimate rotation depends on the number of bolts as well as the diameter of bolts. The rotational ductility calculated by (1) and the energy dissipated are presented in Table. III. There is considerable energy dissipation capacity for all the connections with top and seat angles. The energy dissipation is also higher for angle with more thickness (t) and large number of bolts (specimen S3).

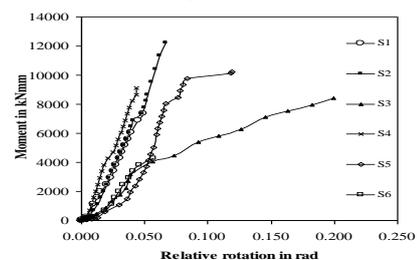


Fig. 7. Moment- relative rotation curve of specimens with Angle Sections

Table. II Energy dissipation for different d/t ratios

Specimen			Number of bolt	Energy dissipated (kNmm rad)
S1	6mm thick	$d/t = 1$	2	177.4
S3	6mm thick	$d/t = 1$	4	1019.8
S4	3mm thick	$d/t = 1$	4	195.4
S6	3mm thick	$d/t = 1$	2	132.4

From Table II, on comparing joints with connection specimens S2 and S3, it is clear that the increase in the ductility is more if the bolt area is changed by increasing the number of bolts rather than increasing the diameter of individual bolt. It is clear from Table III that the ductility can be improved by increasing the number of bolts as well as the angle thickness.

Table. III. Energy dissipated by the connection

Specimen	Ultimate Rotation ϕ_u (rad)	Yield Rotation ϕ_y (rad)	Rotational Ductility $\mu_r = \phi_u / \phi_y$	Energy Dissipated (kNmm.rad)
S1	0.0488	0.0151	3.23	167.72
S2	0.0671	0.0413	1.62	370.18
S3	0.1995	0.0382	5.22	1019.8
S4	0.0434	0.0175	2.48	185.45
S5	0.1192	0.0668	1.78	653.57
S6	0.0572	0.0239	2.39	114.90
S7				208.23

The moment-relative rotation curve under static half cycle loading of double web angle connection (specimen S7) in comparison with top and bottom seat angle connection (specimen S5) is depicted in Fig. 8.

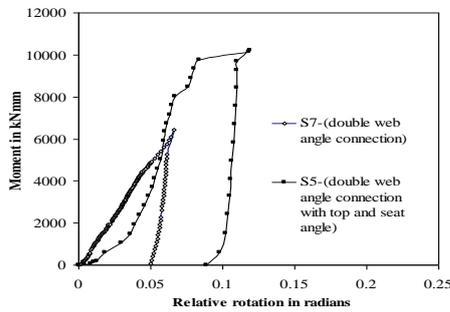


Fig. 8. Moment-relative rotation curve under

From Fig. 8, it is evident that the energy absorption capacity of the connection was improved by using top and seat angles. The energy absorption was increased by 180% by providing seat angles in addition to web angles for moment resistant connection under study. On testing the double web angle connection specimen S7, it was found that the bolts connecting the web angle were bent and the gap was closed. But for the specimen S5 with top and seat angle there was yielding of seat angle followed by breaking of bolt in the seat angle with relatively less bending of the bolts connecting the web angles. The leg of seat angle connecting to the column flange was yielded as shown in Fig. 9.

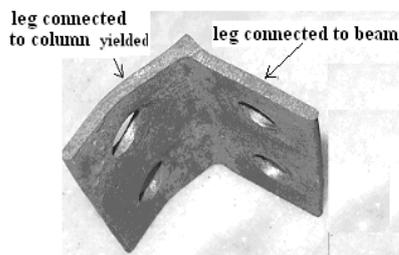


Fig. 9 Deformed seat angle

Finite Element Analysis using ANSYS10

Finite element model of one of the specimen S1 is shown in Fig. 10. In the nonlinear analysis, a maximum load of 30kN was taken and the solution was converged at the ninth sub step (corresponding load was 27kN). The deflected shape of the beam-column joint is shown in Fig. 11. The load-deflection and moment-rotation curves for specimen S1 obtained analytically and experimentally were compared in Fig. 12. Deflection at free end of beam is designated as dl in the graph.

For the specimen S2 the solutions get converged to a load of 39.5kN (14th substep). The moment-rotation curve for specimen S2 obtained analytically and experimentally was plotted as shown in Fig. 13. Similar behaviour was observed in the elastic range without much variation, but there was a deviation of less than 10% of analytical result from the experimental values was seen in the inelastic region. A clear indication of slip at bolt was observed near yield point during experiment and the effect was not simulated analytically. This has resulted in the deviation of analytical result from experimental result in post elastic region. The load-deflection curve for specimen S7 obtained analytically and experimentally was compared as shown in Fig. 14. For the specimen S6 the solutions get converged to a load of 15kN. The moment-rotation curve for specimen S5 obtained analytically and experimentally was plotted as shown in Fig. 15.

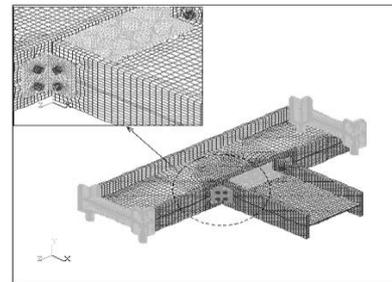


Fig. 10. Model of specimen S1

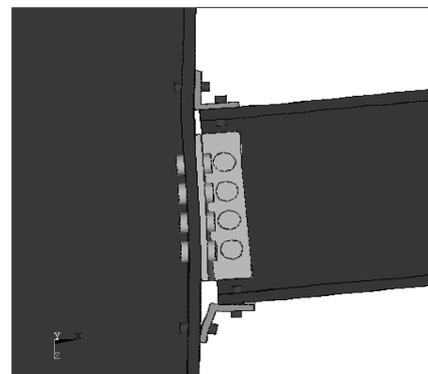


Fig. 11 Deformed model

There was only less than 10 % variation in the analytical result when compared with experimental result. On comparing the analytical and experimental results it can be seen that for same load (or moment), the deflection (or relative rotation) is more in the experimental result. This may be due to the slip in the bolt.

VII. CONCLUSION

Based on the experimental and analytical investigation conducted on bolted steel beam-column joint, the following conclusions are drawn.

- The use of top and seat angles improved the energy dissipation capacity. The energy dissipation of the connection increases as the thickness of top and seat angles and number of bolts increases.
- Increasing the number of bolts rather than increasing the diameter of bolt in the connection with top and bottom seat angles improves ductility.

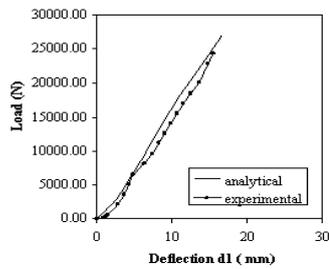
The maximum difference observed between the analytical and experimental results is less than 10%.

NOTATIONS

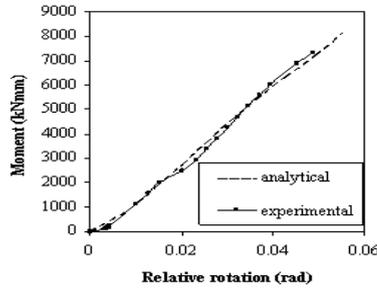
- E - Young's modulus of elasticity of the material
- F - Connection stiffness of beam
- I_C - Moment of inertia of column section.
- I_b - Moment of inertia of the beam cross section.
- L_C - Length of column member
- M - Moment
- Φ - Rotation of the column at the joint
- θ - Rotation of beam at the joint
- Φ_u - Rotation on ultimate state
- Φ_y - Rotation on reaching first yield point of joints.
- d_1 - Deflection at free end of beam.

REFERENCES

- [1] Cheol, M.Y., and Young, M.K., "Cyclic Behaviour of Bolted and Welded Beam to Column Joints", International Journal of Mechanical Sciences, doi:10.1016/j.ijmecsci.2006.09.022.
- [2] Chen, S.J., Yeh, C.H and Chu, J.M., "Ductile Steel Beam-to-Column Connections for Seismic Resistance", Journal of Structural Engineering, ASCE, Vol 122(11), 1996, pp 1292-1299.
- [3] Secer, M. and A. U. Ozturk, "An Investigation for Semi-Rigid Frames by Different Connection Models", Mathematical and Computational Applications, Vol 10, No 1, 2005, pp 35-44.
- [4] Christopoulos, C., Filiatrault, A., Uang, C.M., and Folz, B., "Post Tensioned Energy Dissipating Connections for Moment-Resisting Steel Frames", Journal of Structural Engineering, Vol 128(9), 2002, pp 1111-1120.
- [5] Frederick, G.R., and Abuyasein, O.A., "Analysis of Frames with Semi-Rigid Joints", Computers and Structures, Vol 52, No 6, 1994, pp 1161-1168.
- [6] Kartal M.E., Basaga H.B., Bayraktar A. and Muvafik M., "Effects of Semi-Rigid Connection on Structural Response", Electronic Journal of Structural Engineering, eJSE, Vol 10, 2010, pp 22-35.



a) Load-deflection curve



b) Moment-relative rotation curve

Fig. 12 Specimen S1-6mm thick angle, 2nos 6mm bolt

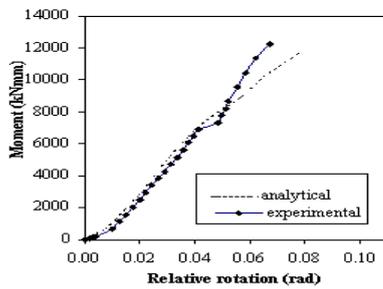


Fig. 13. Specimen S2-6mm thick angle, 2nos 10mm bolt

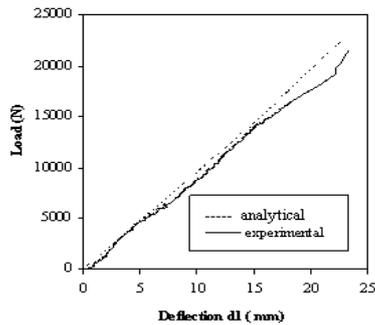


Fig. 14 Specimen S7-double web angle connection

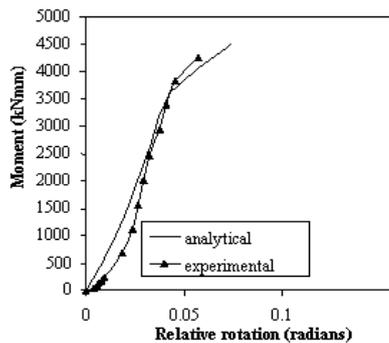


Fig. 15 Specimen S6-3mm thick angle, 2nos 6mm bolt

Mitigation of Harmonics in Distribution System Using SAPF

G. Vamsi Krishna¹, P. Ramesh²

¹M.Tech Scholar, Power Electronics, Nova College Of Engineering & Technology, Jangareddy Gudem, Westgodavari(D.t), AP, India.

²Associate professor, Dept. of Electrical & Electronics Engg. RISE Prakasam group of institutions, Ongole, Prakasam(D.t), AP, India.

ABSTRACT: The typical loads in the commercial buildings comprise of many single phase and three phase loads which includes non-linear loads. These loads are highly sensitive for any input variations in voltage. The performance of the electrical equipment gets worsened if they are supplied with polluted or distorted voltage. This paper presents modeling and analysis of custom power devices. SAPF based UPQC to take power quality problems. Series active power filter is implemented for harmonic and voltage distribution compensation of the three phase three wire distribution system. The custom power devices are realized by the hysteresis control based SAPF. The SAPF can be realized by three phase three leg VSI. A dynamic model of SAPF is developed in the MATLAB/SIMULINK environment and the simulation results are presented to show the effectiveness of proposed SAPF custom power device for a three phase three wire distribution system.

Keywords: Custom Power Devices, Hysteresis Based Controller, MATLAB, SAPF (Series Active Power Filter), SIMULINK, UPQC (Unified Power Quality Conditioner), VSI(Voltage Source Inverter).

I. INTRODUCTION

Power quality phenomena include all possible situations in which the waveform of the supply voltage (voltage quality) or load current (current quality) deviate from the sinusoidal waveform at rated frequency with amplitude corresponding to the rated rms value for all three phases of a three-phase system[1]. The wide range of power quality disturbances covers sudden, short duration variations, e.g. impulsive and oscillatory transients, voltage sags, short interruptions, as well as steady state deviations, such as harmonics and flicker. One can also distinguish, based on the cause, between disturbances related to the quality of the supply voltage and those related to the quality of the current taken by the load [2]. To the first class covers voltage dips and interruptions, mostly caused by faults in the power system. These disturbances may cause tripping of "sensitive" electronic equipment with disastrous consequences in industrial plants where tripping of critical equipment can bear the stoppage of the whole production with high costs associated. One can say that in this case it is the source that disturbs the load. To avoid consistent money losses, industrial customers often decide to install mitigation equipment to protect their plants from such disturbances. The second class covers phenomena due to low quality of the current drawn by the load. In this case, it is the load that disturbs the source. A typical example is current harmonics drawn by disturbing loads like diode rectifiers, or unbalanced currents drawn by unbalanced any direct production loss related to the occurrence of these

power quality phenomena. But poor quality of the current taken by many customers together will ultimately result in low quality of the power delivered to other customers[3]. Both harmonics and unbalanced currents ultimately cause distortion and respectively, unbalance in the voltage as well. Therefore, proper standards are issued to limit the quantity of harmonic currents, unbalance and/or flicker that a load may introduce. To comply with limits set by standards, customers often have to install mitigation equipment. In recent years, both industrial and commercial customers of utilities have reported a rising tide of misadventures related to power quality. The trouble stems from the increased refinement of today's automated equipment, whether variable speed drives or robots, automated production lines or machine tools, programmable logic controllers or power supplies in computers. They and their like are far vulnerable to disturbances on the

utility system than were the previous generation of electromechanical equipment and the previous less automated production and information systems. A growing number of loads is sensitive to customers' critical processes which have costly consequences if disturbed by either poor power quality or power interruption. For the reasons described above, there is a growing interest in equipment for mitigation of power quality disturbances, especially in newer devices based on power electronics called "custom power devices" able to deliver customized solutions to power quality problems. The term Custom Power describes the value-added power that electric utilities and other service providers will offer their customers in the future. The improved level of reliability of this power, in terms of reduced interruptions and less variation, will stem from an integrated solution to present problems, of which a prominent feature will be the application of power electronic controllers to the utility distribution systems and/or at the supply and of many industrial and commercial customers and industrial parks. The compensating devices are used for active filtering; load balancing, power factor correction and voltage regulation.

The active power filters, which eliminate the harmonics, can be connected in both shunt and series. Shunt active power filter can perform power factor correction, harmonic filtering when connected at the load terminals. The harmonic filtering approach is based on the principle of injecting harmonic current into the AC system, of the same amplitude and reverse phase to that of the load current harmonics. Dynamic Voltage Restorer (DVR) is a series connected device. The main purpose of this device is to protect sensitive loads from sag/swell interruptions in the supply side. This is accomplished by rapid series voltage injection to compensate for the drop/rise in the supply voltage. Since this is a series device, it can also be used as a series active power filter. Unified Power Quality

Conditioner (UPQC) [5] is a very versatile device that can inject current in shunt and voltage in series simultaneously in a dual control mode. Therefore it can perform both the functions of load compensation and voltage control at the same time.

II. Sources of harmonics

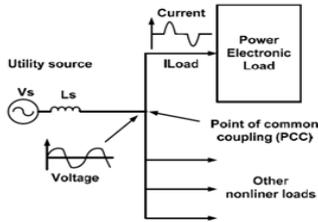


Fig.1. Harmonic problems at source side

In Fig.1. The voltage waveform at the Point Common Coupling (PCC) is distorted. Harmonics are considered as one of the most critical problems in electric power systems. Harmonics in power distribution system are current or voltage that are integer multiples of fundamental frequency. For example if the fundamental frequency is 50Hz, then 3rd is 150Hz, 5th is 250Hz. Ideally, voltage and current waveforms are perfect sinusoids. However, because of the increased popularity of electronic and non linear loads, these waveforms become distorted. This deviation from a perfect sine wave can be represented by harmonic components having a frequency that is an integral multiple of the fundamental frequency. Thus, a pure voltage or current sine wave has no distortion and no harmonics and a non sinusoidal wave has distortion and harmonics. In order to quantify the distortion, the term of Total Harmonics Distortion (THD) is used. The THD value is the effective value of all the harmonics current added together compared with the value of the fundamental current. The simple block diagram in Fig.1. illustrates the distortion problem due to harmonic at low voltage levels.

3.1. Series active power filter

Series active filter is to be placed in series between the ac source and the load (or harmonic source) to force the source current to become sinusoidal. The approach therefore blocking harmonic current flow from the load to the ac source and from the ac source to the load side. The main advantage of series filters over parallel ones is that they are ideal

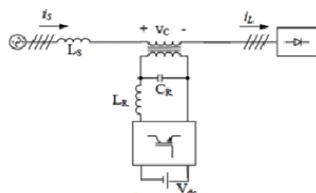


Fig.2. Series active filter

Fig.2. Schematic diagram of Series active filter for eliminating voltage-waveform harmonics, and for balancing three-phase voltages. This, in fact, means that this category of filter is used to improve the quality of the system voltage for the benefit of the load. It provides the load with a pure

sinusoidal waveform, which is important for voltage-sensitive devices.

3.2. Unified Power Quality Conditioner

The UPQC consists of two three phase inverters connected in cascade in such a manner that inverter II is connected in parallel with the load. Inverter I is connected in series with the supply voltage through a transformer. The main purpose of the shunt compensator is to compensate for the reactive power demanded by the load, to eliminate the harmonics and to regulate the common dc link voltage. Fig.3. Schematic diagram of UPQC. The series compensator is operated in PWM voltage controlled mode. It injects voltage in quadrature advance to the supply voltage (current) such that the load end voltage is always maintained at the desired value. The two inverters operate in a coordinated manner.

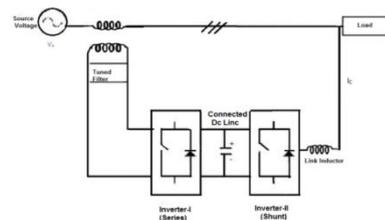


Fig.3. Schematic diagram of UPQC

3.3 Control Scheme of Series Active Power Filter

A simple algorithm is developed to control the series and shunt filters. The series filter is controlled such that it injects voltages (V_{ca}, V_{cb}, V_{cc}) which cancel out the distortions and/or unbalance present in the supply voltages (V_{sa}, V_{sb}, V_{sc}) thus making the voltages at the PCC (V_{ls}, V_{lb}, V_{lc}) perfectly balanced and sinusoidal with the desired amplitude. In other words, the sum of the supply voltage and the injected series filter voltage makes the desired voltage at the load terminals. The control strategy for the series APF is shown in

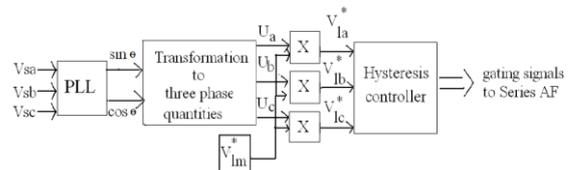


Fig.4. Control scheme of series APF

3.4 Reference Voltage Generation & Hysteresis Voltage Controller

Since the supply voltage is unbalanced and or distorted, a phase locked loop (PLL) is used to achieve synchronization with the supply. This PLL converts the distorted input voltage into pure three phase sinusoidal supply of RMS value of each phase equal to that of the fundamental (1 p.u). Three phase distorted/unbalanced supply voltages are sensed and given to the PLL which generates two quadrature unit vectors ($\sin\theta, \cos\theta$). The sensed supply voltage is multiplied with a suitable value of gain before being given as an input to the PLL. The in-phase sine and cosine outputs from the PLL are used to compute the supply in phase, 120° displaced three unit vectors

(u_a, u_b, u_c) using eqn.(1) as

$$\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix} \begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \times \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix} \quad (1)$$

The computed three in-phase unit vectors are then multiplied with the desired peak value of the PCC phase voltage (V_{lm}^*), which becomes the three-phase reference PCC voltages as

$$\begin{pmatrix} V_{1a}^* \\ V_{1b}^* \\ V_{1c}^* \end{pmatrix} = V_{lm}^* \begin{pmatrix} u_a \\ u_b \\ u_c \end{pmatrix} \quad (2)$$

The desired peak value of the PCC phase voltage is considered to be $338V = \frac{(415 \times \sqrt{2})}{\sqrt{3}}$

The output of the hysteresis controller is switching signals to the six switches of the VSI of the series AF. The hysteresis controller generates the switching signals such that the voltage at the PCC becomes the desired sinusoidal reference voltage. Therefore, the injected voltage across the series transformer through the ripple filter.

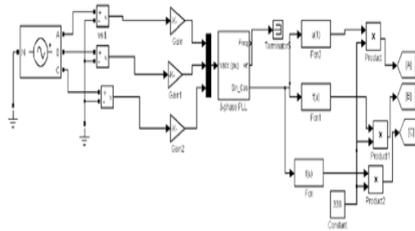


Fig.5. reference voltage generator

cancel out the harmonics and unbalance present in the supply voltage. The MATLAB/Simulink model of the control scheme for series active filter is shown in Fig.5.

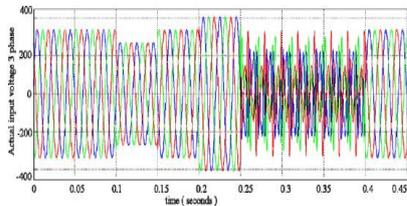


Fig.6. input voltage with distortions

The reference voltages are generated by the Reference voltage generator block is shown in Fig.5.

The system voltage which is fully distorted is shown in Fig.6 After comparing the reference voltages with the load voltage the error signal is passed through Relay block.

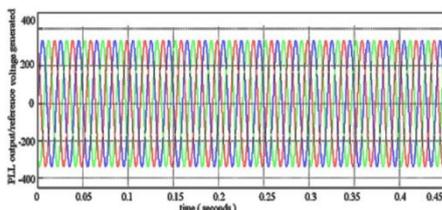


Fig.7. reference voltage generate

Based on comparing the reference voltages with load voltages in a voltage hysteresis controller (Fig.6)

gating pulses are generated and given to IGBTs to compensate the disturbances in the system.

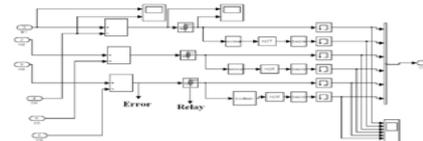


Fig.8.hysteresis voltage controller

The main advantage here to use a hysteresis band controller over a P-I controller is that the former does not require the specifications regarding the system parameters i.e switching frequency , load angle etc. but the latter controller (P-I) requires an additional design criteria for its application analyzed by taking the FFT analysis of the Source and load Voltages.

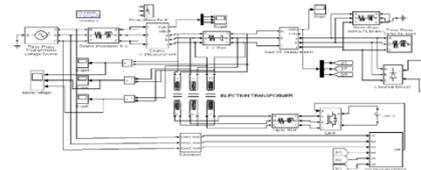


Fig.9. MATLAB/Simulink model of proposed series active power filter

III. SIMULATION RESULTS

Case 1

Fig.10 shows the source voltage in which rated 1 p.u voltage is created from 0 to 0.1 seconds , 0.8 p.u sag from 0.1 to 0.15 seconds, 1 p.u voltage from 0.15 to 0.2 seconds, 1.2 p.u swell from 0.2 to 0.25 seconds , 0.4 p.u sag from 0.25 to 0.3 seconds , 0.9 sag from 0.3 to 0.4 seconds and 1p.u voltage from 0.4 to 0.5 seconds. R=50Ω and L=1 mH The SAPF is simulated with only interruptons in the form of Sag, Swell in the input side & the performance of SAPF is

Case – I

Time(seconds)	Voltage (p.u)
0 to 0.1	1
0.1 to 0.15	0.8
0.15 to 0.2	1
0.2 to 0.25	1.2
0.25 to 0.4	(5 th & 7 th order harmonics)
0.4 to 0.5	0.9

Table.1. Source Voltage 3-phase (line- Ground)

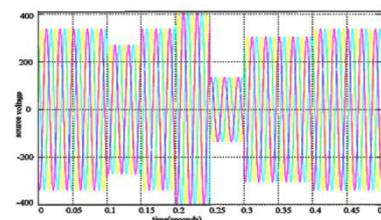


Fig.10.source voltage

Fig.11. shows the compensated voltage injected by each phases to cancel the source side disturbances present in the system.

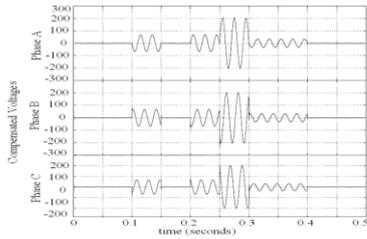


Fig.11. compensated voltages injected for each phases

Due to the injection of the above voltages through the injection transformer in series with the line the load voltage is sinusoidal as shown in the Fig.12.

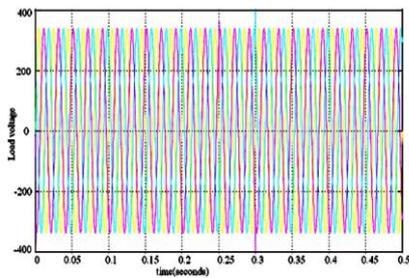


Fig.12. load voltage

The Total harmonic distortion of source voltage is 1.64% and load voltage is 0.04 % as shown in Figs 13,14 respectively.

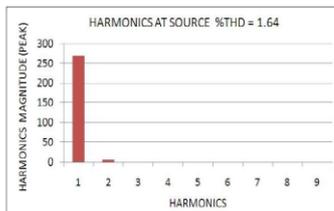


Fig.13. source voltage harmonic spectrum

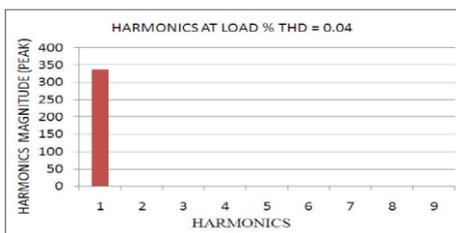


Fig.14. load voltage harmonic spectrum

The SAPF is simulated with only interruptions in the form of Sag, Swell in the input side & the performance of SAPF is analyzed by taking the FFT analysis of the Source and load Voltages.

Case .2

Fig.15. shows the source voltage in which rated 1 p.u voltage is created from 0 to 0.1 seconds , 0.8 p.u sag from 0.1 to 0.15 seconds, 1 p.u voltage from 0.15 to 0.2 seconds, 1.2 p.u swell from 0.2 to 0.25 seconds , 5th order harmonics of 0.4 p.u and 7th order harmonics of 0.2 p.u from 0.25 to 0.4 seconds, 0.9 sag from 0.4 to 0.5 seconds . the load is R=50Ω and L=10 mH.

Case-II

Time(seconds)	Voltage (p.u)
0 to 0.1	1
0.1 to 0.15	0.8
0.15 to 0.2	1
0.2 to 0.25	1.2
0.25 to 0.3	0.4
0.3 to 0.4	0.9
0.4 to 0.5	1

Table.2. Source Voltage 3-phase (line-Ground)

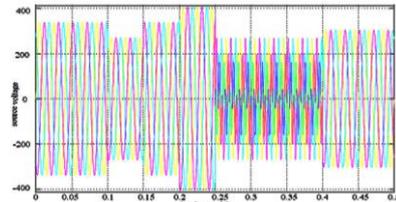


Fig.15. source voltage

Fig.16. shows the compensated voltage injected by each phases to cancel the source side disturbances present in the system.

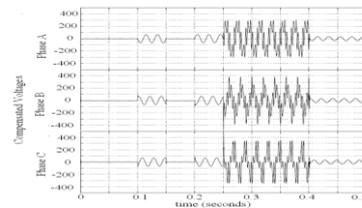


Fig.16. Compensated voltages injected for each phases

Due to the injection of the above voltages through the injection transformer in series with the line the load voltage is sinusoidal as shown in the Fig. 17.

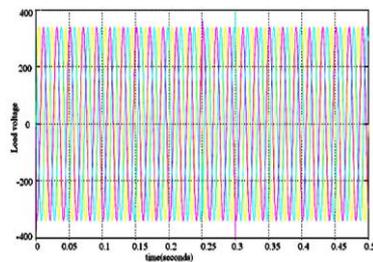


Fig.17. load voltage

The Total harmonic distortion of source voltage is 64.37% and load voltage is 0.72 % as shown in Fig.18, Fig.19. respectively.

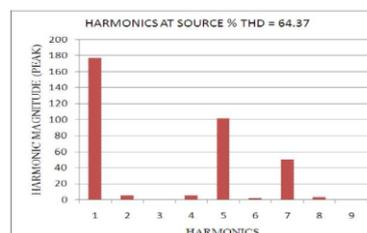


Fig.18. source voltage harmonic spectrum

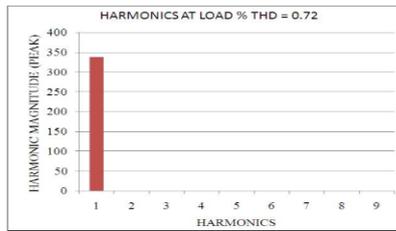


Fig.19.load voltage harmonic spectrum

The SAPF is simulated with interruptions in the form of Sag, Swell & harmonics in the input side(5th & 7th) & the performance of SAPF is analyzed by taking the FFT analysis of the Source and load Voltages.

IV. Conclusion

A Series active power filter has been investigated for power quality improvement. Various simulations are carried out to analyze the performance of the system. Hysteresis controller based Series active power filter is implemented for harmonic and voltage distortion compensation of the non-linear load. The Simulation is even extended for abnormal faults occurring on the power system like L-G & L-L faults. The simulation results of series active power filter has shown the ability to compensate voltage sag, swell and harmonics present at input source side. With these functions, the proposed SAPF is suitable for connecting at the PCC of industrial drives which are most sensitive to Sags, Swells and harmonics. The THD of the load voltage is below 5%, the harmonics limit imposed by IEEE standard.

REFERENCES

- [1] Bhim Singh and Venkateswarlu.P, "A Simplified Control Algorithm for Three- Phase, Four-Wire Unified Power Quality Conditioner" Journal of Power Electronics, Vol. 10, No. 1, January 2010 pp.91-96.
- [2] B.Singh,K.Al-Haddad,and A.Chandra, "A review of active power filters for power quality improvement,"IEEE Trans.Ind.Electron., vol.45,no.5,pp.960-971,Oct 1999.
- [3] V.Khadkikar,A.Chandra,A.O.Barry,and T.D.Nguyen, "Application of UPQC to protect a sensitive load on a polluted distribution network," in Proc.IEEE PES General Meeting, Montreal,QC,Canada,2006,6pp.
- [4] V.Khadkikar,A.Chandra,A.O.Barry,and T.D.Nguyen, "Conceptual analysis of unified power quality conditioner (UPQC)," in proc.IEEE ISIE,2006,pp.1088- 1093.
- [5] Vinod Khadkikar and Ambrish Chandra, "A Novel Structure for Three-Phase Four-Wire Distribution System Utilizing Unified Power Quality Conditioner (UPQC)" IEEE Transactions on Industry Appl., Vol. 45, No. 5, September/October 2009.
- [6] S.K. Jain, P. Agrawal and H.O. Gupta, "Fuzzy logic controlled shunt active power filter for power quality improvement"IEEE Proceedings on Elect.Power.Appl.Vol 149.No.5,September 2002.
- [7] Ash Pal, A. Swarup, and Bhim Singh, " A Comparison of Three Topologies of Three-Phase Four-Wire UPQC for Power Quality Improvement" 16th National Power systems conference, 15th-17th December, 2010 pp.227-232.
- [8] S. Chakraborty and F. A. Farret, " A Single Phase High Frequency AC Microgrid with an Unified Power Quality Conditioner," Proceedings on IEEE 2003, pp.952-967.
- [9] Active power filter with automatic control circuit for neutral current harmonic minimization technique by Mohd Izhar Bin Bakar Handbook of Power Quality by Angelo Bagini

BIOGRAPHICS

G. Vamsi Krishna received B.Tech degree in Electrical and Electronics Engineering from Dvr & Dr Hs MIC College of Technology affiliated to JNTUH , India in the year 2008.He is presently M.Tech scholar in Power Electronics at Nova college of Engineering & Technology affiliated to JNTU Kakinada, A.P, India.

P. Ramesh received M.Tech from JNTU Hyderabad, AP,India. He is doing his PhD at Dr MGR University Chennai. He has published many journals in national and international level. He also Associative member of IEEE and Life member of ISTE.His area of research is soft computing, power systems.

Experimental Investigation of Performance Parameter of Diesel Engine Operating On Methyl Tertiary Butyl Ether (MTBE)

Mukesh kumar Bunkar¹, Nitin Shrivastava², Vipin Shrivastava³
 Dept. of Mech. Engg, University Institute of Technology, RGPV- Bhopal, India

Abstract: This study investigated the performance parameters of direct injection diesel engine using methyl tertiary butyl ether (MTBE) blended with diesel. The entire investigation performance of diesel engine is carried in two phases using the 100% diesel in the first phases and the MTBE blended with diesel fuels in the second phase. The MTBE blended diesel fuels are prepared in the different proportions of 5 % & 10 % MTBE blends. The experimental performance consists of a diesel engine coupled with rope brake dynamometer. All the experiments are performed at a constant speed of 1500 rpm by varying the load and then data obtained from the experiments are used to evaluate the performance parameters of diesel engine. It is investigated that the brake thermal efficiency (BTE) of diesel engine using MTBE blends were decreased as the amount of blends increased in diesel. The BTE of diesel engine using 5% and 10% blends are 25.26 % & 24.65 % respectively that is less from 26.35% than that of diesel. The brake specific energy consumption of 5 & 10 percent of MTBE blends consumed higher energy in comparison to diesel fuel during tested on diesel engine. It is investigated that the brake specific fuel consumption of diesel engine using MTBE blends are increased for all brake load than that of diesel fuel.

Keywords: MTBE, Diesel Engine, Performance.

I. Introduction

Diesel engines have been commonly used in recent decades as an alternative power source for light or heavy duty vehicles because of the economical and environmental reason. Therefore the worldwide diesel fuel consumption has increased with the populations sourced from diesel engine [1, 2]. In normally, the performance enhancement of the diesel engines are prominently achieved by three methods such as engine modification, fuel adulteration and exhaust gas treatment. Most of the researchers have put their valiant efforts on fuel adulteration techniques as they do not require any major hardware modification. The effect of fuel adulteration of the conventional fossil fuel (diesel) has been materialized on incorporating some metallic additives [3, 4], oxygenated additives [5] and ignition enhancement additive [6]. The oxygenated additives such as Dimethyl Carbonate (DMC), ethylene glycol mono-acetate, 2-Methoxyethyl acetate, Diethyl Ether (DEE), Methyl Tetra Butyl Ether (MTBE), Dimethyl Ether (DME), have been blended with diesel fuel in various proportions and achieved better performance and emission characteristic in a diesel engine. The combustion pressures and temperatures of the 5 vol % MTBE blend are similar to or a very little lower than that of neat diesel fuel. However, the 10 and 15 volume % MTBE blends show considerable decreases in the combustion pressures and temperatures [7].

The emissions such as unburned Hydrocarbon and CO emissions are also reduced for the MTBE blended diesel fuels due to the shortened ignition delay and complete combustion compared to that of neat diesel operation. Sivakumar et al. observed the lower blends of MTBE and found that MTBE can be considered as a potential diesel additive for the enhancement of brake thermal efficiency and considerable reduction of unburned Hydrocarbon, CO and smoke in diesel engine [8].

II. Experimental Setup and Methodology

Experimental set up is shown in Figure 1



Fig 1 Experimental setup

The Specifications of the experimental setup are given in Table 1

Table.1 Technical Specification of test engine	
Company and Model	Kirloskar oil Engine , SV1
Type	Single cylinder, 4- Stroke, diesel engine
Bore	87.5mm
Stroke	110mm
Rpm	1500rpm
Rated power	8 HP
Type of cooling	Water cooled
Compression ratio	16.5:1

The fuel properties are shown in Table 2

Table.2 Properties of diesel and biodiesel		
Properties	Diesel	MTBE
Specific gravity (gm/cm ³)	0.836	.7404
Calorific value (kJ/kg)	42850	35108
Cetane number	48	-
Density(kg/l),(g/cm ³)	.836	.7405
Boiling point (°C)	154.30	55.2

Melting Point(°C)	-30 -18	-108°
Chemical formula	C ₁₄ H ₂₂	C ₅ H ₁₂ O
Flash point (°C)	56	-28.2
Fire point (°C)	64	480
Carbon (%)	86	-
Hydrogen (%)	14	-

The MTBE is blended with diesel fuel with the proportions of 5 % and 10 % by volume using a measuring cylinder. The blends of MTBE and diesel fuel are subjected to diesel engine and find that the performance of diesel engine using these blends. The performance parameters studied are fuel consumption, Brake specific fuel consumption, Brake specific energy consumption, and Brake thermal efficiency.

III. RESULT AND DISCUSSION

3.1 Fuel Consumption (FC)

Figure 2 shows the Variation of fuel consumption for neat Diesel, 5 and 10 percent blends of MTBE. It was resulted that the fuel consumption of diesel fuel is higher when the MTBE blends are used 5% and 10%. As load is increased the FC is slightly increased in comparison to neat diesel. The FC of 5% and 10 % consumed 5% and 10% more fuel as compared to conventional diesel fuel.

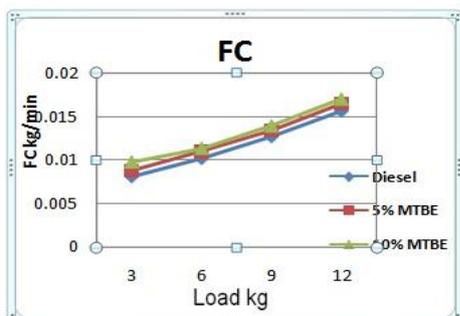


Figure 2. Variation of fuel consumption with load

3.2 Brake Specific Fuel Consumption (BSFC)

Figure 3 shows the Variation of brake specific fuel consumption for Diesel, 5% & 10% MTBE blends. The brake specific fuel consumption is an essential parameter to compare engines and determine of fuel efficiency of an engine. It was resulted that the brake specific fuel consumption (BSFC) is higher than the diesel fuel when the MTBE blends are used 5% and 10%. The BSFC slightly decreased with increased the load. The BSFC of 5% and 10 % consumed approximately 4% and 7% more fuel than that of diesel fuel.

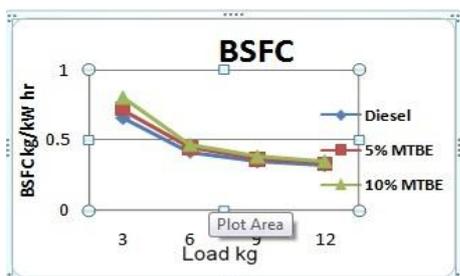


Figure 3. Variation of Brake Specific fuel consumption with load

3.3 Brake Specific Energy Consumption (BSEC)

Figure 4 shows the Variation of brake specific energy consumption for Diesel, -5% & 10% MTBE blends. It was resulted that the brake specific energy consumption is higher than that of diesel fuel when the MTBE blends are used 5% and 10%. The BSEC slightly decreased with increased the load. The BSEC of 5% and 10 % consumed 5.34% and 9% more as compared to diesel fuel.

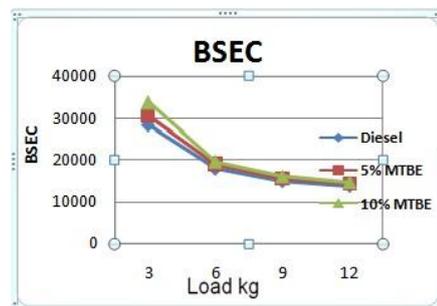


Figure 4. Variation of Brake specific Energy Consumption with load

3.4 Brake Thermal efficiency (BTE)

Figure 5 shows the variation of BTE (Brake thermal efficiency) for 5% and 10% MTBE blends. It is investigated that when the 5% and 10% MTBE blends are used in diesel engine decreased brake thermal efficiency approximately 4-7% than that of diesel fuel. The result showed of 10% MTBE blend the poor thermal efficiency in comparison to neat diesel and 5 % blends.

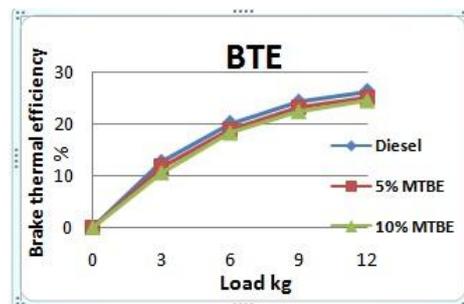


Figure 5. Variation of Brake Thermal Energy with load

IV. Conclusion

The performance of MTBE blended diesel fuel are evaluated for signal cylinder direct injection diesel engine. It is concluded that the brake thermal efficiency of MTBE blends are is decreased by varying load and brake specific energy consumption of diesel engine are increased as the blends are increased in diesel.

References

- [1] A.C. Lloyd, T.A. Cackette, "Diesel engines: environmental impact and control", Journal of the Air & Waste Management Association 51 (2001) 809e847.
- [2] A.N. Shah, G. Yun-shan, T. Jian-wei, Carbonyls emission comparison of a turbocharged diesel engine fuelled with diesel, biodiesel, and biodieselediesel blend, Jordan Journal of Mechanical and Industrial Engineering 3 (2) (2009) 111e118.
- [3] Arul Mozhi Selvan, V., Anand R. B., and Udayakumar M. "Effects of cerium oxide nanoparticle addition in Diesel and diesel-biodiesel-ethanol blends on the performance and

- emission characteristics of a CI engine” Journal of Engineering and Applied Sciences, 2009, 1-6.
- [4] Sadhik Basha, J., and Anand, R.B., “ Application of nanoparticle/nanofluid in compression ignition engines – A case study” International journal of applied engineering and research 2010, 697-708.
- [5] Higgins, B., Siebers, D., Mueller C., and Aradi, A., “Effects of an ignition-enhancing, diesel-fuel additive on diesel-spray evaporation, mixing, ignition, and combustion” Twenty-Seventh Symposium (International) on Combustion/The Combustion Institute, 1998, 1873–1880.
- [6] Nubia M. Ribeiro, Angelo C. Pinto, Cristina M. Quintella, Gisele O. da Rocha, Leonardo S. G. Teixeira, Lílian L. N. Guarieiro, Maria do Carmo Rangel, Márcia C. C. Veloso, Michelle J. C. Rezende, Rosenira Serpa da Cruz, Ana Maria de Oliveira, Ednildo A. Torres, and Jailson B. de Andrade. “The role of additives for diesel and diesel blended (Ethanol or Biodiesel) fuels: a review” Energy & Fuels 2007, 21, 2433-2445.
- [7] Roy, M.M., “Investigation of methyl tertiary butyl ether–diesel combustion and odorous emissions in a direct-injection diesel engine”, Proc. IMechE, Part D, J. Automobile Engineering, 222, 2008, 251-26
- [8] V. Sivakumar, J.Sarangan, R. B. Anand, “Performance, Combustion and emission Characteristics of a CI Engine using MTBE Blended Diesel Fuel, IEEE, 2010, 170-174.

Implementation of Taguchi methodology to Optimization of CNC end milling process parameters of AL6351 –T6

Pinki Maurya¹, Dr. Pankaj Sharma², Bijendra Diwaker³

Department of Mechanical Engineering

* ** (Rajasthan Institute Of Engineering & Technology (RIET), Jaipur, India)

*** (Institute of Technology & Management (ITM) University, Gwalior, India)

Abstract: In this paper we have study on CNC end milling, influence of various machining parameters like, tool feed (mm/min), tool speed (rpm), tool diameter (mm) and depth of cut (mm). In the present study, experiments are conducted on AL 6351 –T6 material with three levels and four factors to optimize process parameter and surface roughness. An L9 (3*4) Taguchi standard orthogonal array (OA) is chosen for design of experiments and the main influencing factor are determined for each given machining criteria by using Analysis of variance (ANOVA). The surface finish have been identified as quality attributes and are assumed to be directly related to productivity. In this experiment we were found that order of significant of main parameter decreasing order is M3>N2>O2>P1.(Tool feed(M), Tool speed(N), Tool diameter(O) and Depth of cut (P)).

Keywords: CNC end milling, Surface roughness (SR), Taguchi methodology

I. INTRODUCTION

Today, industry needs quality and productivity. The increase of consumer needs for quality metal cutting products has driven the metal cutting industry to continuous improve quality control of metal cutting process. The end milling process is one of the most fundamental processes of metal removing process. In order to obtain better surface roughness, the proper setting of cutting parameters is crucial before the process takes place. Several factors will influence the final surface roughness in a CNC milling operation. The final surface roughness might be considered as the sum of two independent effects: 1.The ideal surface roughness is are sult of the geometry of tool and feed rate.2.The natural surface roughness is a result of the irregularities in the cutting operation Factors such as spindle speed, feed rate, tool diameter and depth of cut that control the chip formations, or the material properties of both tool and work piece are Even in the occurrence of chatter or vibrations of the machine tool, defects in the structure of the work material. In end milling, surface finish and material removal rate are two important aspects, which require attention both from industry personnel as well as in Research & Development, because these two factors greatly influence machining performances. In modern industry, one of the trends is to manufacture low cost, high quality products in short time. Automated and flexible manufacturing systems are employed for that purpose. CNC machines are considered most suitable in flexible manufacturing system. Above all, CNC milling machine is very useful for both its flexibility and versatility. These machines are capable of achieving reasonable accuracy and surface finish. Processing time is also very low as compared to some of the conventional machining process. Which indicates processing time of the work piece, is another important factor that greatly influences production rate and cost? So, there is a need for a tool that should allow the evaluation of the surface roughness before the machining of the part and which, at the same time, can easily be used in the production-floor environment contributing to the minimization of required time and cost and the production of desired surface quality.

II. METHODOLOGY

A. Analysis of Variance (ANOVA):

Analysis of variance (ANOVA) and the F-test (standard analysis) are used to analyse the experimental data [1 ,2 ,3] :

$$CF = T^2/n$$

$$S_T = \sum_{i=1}^{1027} Y_i^2 - CF$$

$$S_x = (Y_{x1}^2/N_{x1} + Y_{x2}^2/N_{x2} + Y_{x3}^2/N_{x3}) - CF$$

$$f_x = (\text{number of levels of parameter } x) - 1$$

$$f_e = f_T - \sum f_x$$

$$f_T = (\text{total number of results}) - 1$$

$$V_x = S_x/f_x$$

$$S_e = S_T - \sum S_x$$

$$F_x = V_x/V_e$$

$$V_e = S_e/f_e$$

$$P_x = S_x'/S_T * 100\%$$

$$S_x' = S_x - (V_e * f_x)$$

$$P_e = (1 - \sum P_x) * 100\%$$

Where;

- CF → correction factor
- n → total number of experiment
- T → total of all results
- S_T → total sum of squares to total variation
- Y_i → value of results of each experiment (i= 1 to 27)
- S_x → sum of squares due to parameter x (x= P, Q, R And S)
- N_{x1}, N_{x2}, N_{x3} → repeating number of each level (1, 2, 3) of parameter x
- Y_{x1}, Y_{x2}, Y_{x3} → value of results of each level (1, 2, 3) of parameter x
- f_x → degree of freedom (DOF) of parameter x
- f_e → degree of freedom (DOF) of error term
- f_T → total degree of freedom
- V_x → variance of parameter x
- S_e → sum of squares of error term
- F_x → F- ratios of parameter x
- V_e → variance of error term
- P_x → percentage contribution of parameter x
- S_x' → pure sum of square
- P_e → percentage contribution of error term

B. Design of Experiment or parameter and level setting

We were design experiment to conducting on CNC end milling of three level and four parameter for the optimization of lower surface roughness through the experimental setup by using Taguchi methodology

TABLE I. CONTROL PARAMETER AND THEIR LIMITS

Control Parameter	Level			Observed Values
	1	2	3	
M: Tool feed (mm/min)	1500	2000	2500	Surface Roughness (SR)
N: Tool speed (rpm)	4000	5000	6000	
O: Tool diameter (mm)	10	12	16	
P: Depth of cut (mm)	0.70	0.80	0.90	

TABLE II L9 ORTHOGONAL ARRAY AND OBSERVED VALUES

No. of trial	Control Parameter(Level)				Result/Observed Value		
	M	N	O	P	Surface Roughness (Ra)		
					1	2	3
1	(1500)1	(4000)1	(10)1	(0.70)1	1.34	0.77	1.18
2	(1500)1	(5000)2	(12)2	(0.80)2	0.9	0.84	1.02
3	(1500)1	(6000)3	(16)3	(0.90)3	1.72	1.12	0.78
4	(2000)2	(4000)1	(12)2	(0.90)3	0.66	0.62	0.67
5	(2000)2	(5000)2	(16)3	(0.70)1	0.74	0.78	0.69
6	(2000)2	(6000)3	(10)1	(0.80)2	0.83	2.06	2.35
7	(2500)3	(4000)1	(16)3	(0.80)2	0.71	0.84	0.64
8	(2500)3	(5000)2	(10)1	(0.90)3	0.81	0.82	0.63
9	(2500)3	(6000)3	(12)2	(0.70)1	0.92	0.66	0.72

TABLE II. ANALYSIS OF VARIANCE AND F-TEST FOR SR

Parameter	DOF	S _x	V _x	F-ratio(F _x)	S _x '	P _x
M	2	0.578585	0.289293	2.563641*	0.352896	7.103917
N	2	1.088807	0.544404	4.824373**	0.863119	17.37486
O	2	0.851163	0.425581	3.7714*	0.625474	12.591
P	2	0.417874	0.208937	1.851549*	0.192185	3.86875
e	18	2.0312	0.112844			

**Significant Parameter, * sub significant parameter

TABLE III. SUMMARIZATION OF SIGNIFICANT PARAMETER ON THE MACHINABILITY OF END MILLING

Parameter	Surface Roughness (SR)
M	*
N	**
O	*
P	*

C. Conformation test

We have found optimal parameter setting for SR as (M3, N2, O2, P1) According to predicted parameter setting with ANOVA calculation. We have conducted the experiments and found SR is **0.69 Ra**. This shows the successful validation of the technique.

spindle horsepower The machine is capable of a three-axis movement (along the x, y, and z planes). CNC programs can be developed in the hyper mill software. The work piece material used was AL6351 –T6 in the form of 420mm x 120mm x 20mm plate. The surface roughness measured by Mitutoyo surface roughness tester

III. Experiments

Experimental Setup

The study was carried out using a BFW V-30 CNC vertical milling machine equipped with a maximum Spindle speed 10000 rpm, feed rate 20 m/min, multiple tool-change capabilities (max number of tools = 25) and with 15 HP

IV. RESULTS AND DISCUSSION

Discusses the result for the CNC end milling process parameter we were found main influence control parameter for the minimizing the surface roughness (SR) is that Tool feed (mm/min). Show table V.

TABLE V RESULTS OF THE CONFIRMATION EXPERIMENT FOR SR

Control Parameter	Level	SR (Ra)
M: Tool feed (mm/min)	3	2500
N: Tool speed (rpm)	2	5000
O: Tool diameter (mm)	2	12
P: Depth of cut (mm)	1	0.7

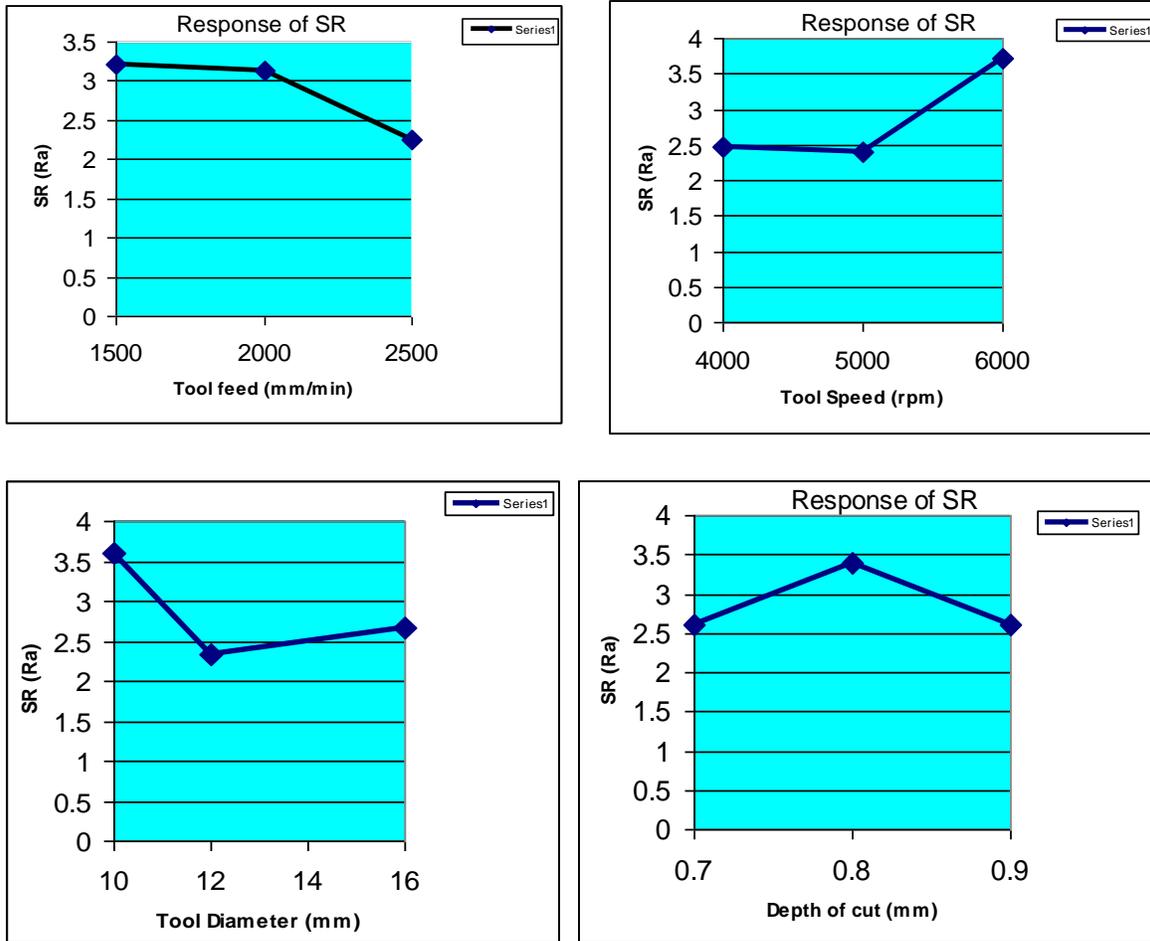


Figure 1: Main effects of each parameter on SR

V. CONCLUSION

In this study the analysis of confirmation experiment and the design of control parameter with three level & four parameters to find the optimal control parameter to minimize the surface roughness that the parameter is **Tool feed (M)** as shown in fig.1 and table V, and the order of significance parameter is **M3>N2>O2>P1**. This is the successful validation of the Taguchi methodology

References

- [1] Yan B.H., Wang C.C., Liu W. D. and Huang F.Y.,(2000) "Machining Characteristics of Al₂O₃/6061 Al Composit using Rotary EDM with a Disklike Electrode" , The International Journal of Advanced Manufacturing Technology, 16 pp.322-333.
- [2] Roy Ranjit K.,(1996) "A Premier On The Taguchi Method" , Van Nostrand Reinhold, New York.
- [3] Roy Ranjit K.,(2001) "Design of Experiments Using the Taguchi Approach" , John Wiley & Sons, New York.
- [4] Routara B.C. & Bandyopadhyay A. & Sahoo P. (2008) "Roughness modeling and optimization in CNC end milling using response surface method effect of workpiece material variation.
- [5] A handbook of a primer on the Taguchi Method by: Ranjit.K.Roy.
- [6] Yang L.J. and Chen C.J., (2001). A systematic approach for identifying optimum surface roughness performance in end-milling operations, *Journal of Industrial Technology*, Vol 17, pp.1-8.
- [7] Ghani J.A., Choudhury I.A. and Hassan H.H., 2004. Application of Taguchi method in the optimization of end milling parameters, *Journal of Material Processing Technology*, Vol. 145, No. 1, pp. 84-92.
- [8] Oktem H., Erzurumlu T. and Kurtaran H., 2005. Application of response surface methodology in the optimization of cutting conditions for surface roughness, *Journal of Material Processing Technology*, Vol. 170, No. 1-2, pp. 11-16.

Impact of Climatological Parameters on Yield of Wheat Using Neural Network Fitting

Parekh, F. P.¹, Suryanarayana, T. M. V.²

^{1,2} (Water Resources Engineering and Management Institute, Faculty of Technology & Engineering, The Maharaja Sayajirao University of Baroda, Samiala, 391410, Gujarat, India)

ABSTRACT : The study was carried out to determine the predominance of various meteorological data on yield of wheat. The meteorological data for the Rabi season are collected and correlated with yield of wheat in Vallabh Vidyanagar for the period 1981- 1999 using neural network fitting. Then, the model is re-trained until the best coefficient of correlation is obtained and this corresponding model is considered as the best model and this is further used to validate the dataset from 2000 to 2006. This whole procedure is repeated for three different Alternatives. In Alternative 1, only maximum and minimum temperatures are correlated with yield data. In Alternative 2, maximum and minimum temperature and relative humidity are correlated with yield data. In Alternative 3, Maximum and minimum temperatures, sunshine hours and relative humidity are correlated with yield. The correlation between maximum and minimum temperatures & relative humidity and yield, the co-efficient of correlation for training, i.e. 1.00 validations, i.e. 0.97 and testing, i.e. 0.95 are best for 70%, whereas in 30% dataset, R comes out to be 0.62, which is good. It can be evidently concluded from the study that considering three variables, the correlation is achieved as the best. This reveals that yield of a crop is very much depended on maximum and minimum temperatures & relative humidity.

Keywords: Climatic variability, Coefficient of correlation, Crop yield, Meteorological data, Neural Fitting.

I. INTRODUCTION

Climatic variability is the major factor influencing the agriculture productivity. Global climate change and its impacts on agriculture have becoming an important issue. Agriculture production is highly dependent on climate and it also adversely affected by increasing climatic variability. The aim is to develop the methodology for assessing this component of the total impact of climate variability on agricultural productivity. There is a need to quantify climatic variability to assess its effect on crop productivity. Among the major wheat growing states in India, Gujarat ranked 6th in production and having productivity of 2,377 kg/ha.

Agrometeorological models are defined as the product of two or more weather factors each representing functioning between yield and weather. These models do not require hypothesis of the plant and environment process. Thus the input requirement is less stringent but the output information is more dependent on the input data. Thus grometeorological models are a practical tool for the analysis of crop response to weather and estimating the yield. Khashei-siuki et al. (2011) studied the ability of Artificial Neural Network (ANN) technology and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for the prediction of dryland wheat (*Triticum aestivum*) yield, based on the

available daily weather and yearly agricultural data. Maqsood et al. (2004) presented the applicability of an ensemble of artificial neural networks (ANNs) and learning paradigms for weather forecasting in southern askatchewan, Canada. Parekh and Suryanarayana (2012) carried out the study to find the quantitative relationship between weather parameters and district level yield of wheat by seasonal trend analysis.

The present study was undertaken with a view to determine the predominance of various meteorological data on yield of wheat in Vallabh Vidhyanagar, Gujarat, India wherein the different weather parameters that are considered will be maximum & minimum temperatures, relative humidity, wind velocity and sunshine hours.

II. STUDY AREA

The entire Gujarat is divided into the various agro-climatic zones. Vallabh Vidyanagar is located in the Anand district and lies in middle Gujarat agro-climatic zone III of Gujarat state. Vallabh Vidyanagar is located at 22°32' N latitude, 72°54' E longitude at an altitude of 34 m above mean sea level. It is bounded on the north by the Kheda district and south by the Gulf of Khambhat, on the west by Ahmedabad district and, on the east by Vadodara district.

The climate of Vallabh Vidyanagar is semi-arid with fairly dry and hot summer. Winter is fairly cold and sets in, in the month of November and continues till the middle of February. Summer is hot and dry which commences from mid of February and ends by the month of June. May is the hottest month with mean maximum temperature around 40.08 °C. The average rainfall is 853 mm.

The soil of the region is popularly known as Goradu soil. It is alluvial in origin. The texture of the soil is sandy loam and black. The soil is deep enough to respond well to anuring and variety of crops of the tropical and sub-tropical regions. The soil is low in organic carbon and nitrogen, medium in available phosphorus and available sulphur. In this area paddy, tur, cotton and ground nut, til are grown in kharif season. In rabi season wheat, gram, and jowar are grown. Especially, in summer season the bajara and ground nut are grown. Tobacco is grown from August and harvested in March. In last few years, there is increase in amount of rainfall which facilitated in agriculture production and various irrigation scheme.

III. DATA COLLECTION

The data required for evaluation in this study are collected from India Meteorological Department, Pune and Krishi Bhavan, Gandhinagar. Long term meteorological daily data are collected from IMD (Indian

Meteorological Department), Pune for Vallabh Vidyanagar, town of Anand district of Gujarat.

The basic weekly meteorological data used comprises of Maximum and minimum temperature($^{\circ}\text{C}$), Relative humidity (%) and Sunshine hours (hours). The yield data of Wheat grown in Vallabh Vidyanagar are collected from the Krishi bhavan, Gandhinagar from year 1981-2006.

IV. METHODOLOGY

The weekly meteorological data viz. temperature, relative humidity, sunshine hours are converted into average monthly data and seasonal data. Wheat is cultivated in Rabi season.

Some data gaps or missing values are indentified in data. The missing values were found out with the SPSS 11.5 software. The linear trend at a point method is used to find the missing values of the data. To study the impact of climatological data on yield of wheat, Neural fitting tool of Artificial Neural Network (ANN) of MATLAB is used.

4.1 Artificial Neural Networks (ANN)

ANN was first introduced as a mathematical aid and was inspired by the neural structure of the brain. An input layer, which is used to present data to the network. An output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors. The ability of a neural network to process information is obtained through a learning process, which is the adaptation of link weights so that the network can produce an approximate output. In general, the learning process of an ANN will reward a correct response of the system to an input by increasing the strength of the current matrix of nodal weights.

There are several features in ANN that distinguish it from the empirical models. First, neural networks have flexible nonlinear function mapping capability which can approximate any continuous measurable function with arbitrarily desired accuracy, whereas most of the commonly used empirical model, do not have this property. Second, being non-parametric and data-driven neural networks impose few prior assumptions on the underlying process from which data are generated. Because of these properties, neural networks are less susceptible to model misspecification than most parametric nonlinear methods.

An ANN can be defined as data processing system consisting large number of simple highly interconnected processing elements (PEs or artificial neurons) in architecture analogous to cerebral cortex of brain. An ANN consists of input, hidden and output layers and each layer includes an array of artificial neurons. A typical neural network is fully connected, which means that there is a connection between each of the neurons in any given layer with each of the neuron in next layer. An artificial neuron is a model whose components are analogous to the components of actual neuron in next layer. An artificial neuron is a model whose components are analogous to the components of actual neuron. The array of input parameters is stored in the input layer and each input variable is represented by a neuron. Each of these inputs is modified by a weight (sometimes called synaptic weight) whose function is analogous to that of the synaptic junction in a biological neuron. The neuron (processing element)

consists of two parts. The first part simply aggregates the weighted inputs resulting in a quantity 1: the second part is essentially a nonlinear filter, usually cooed the transfer function or activation function. The activation function squashes or limits the values of the output of an artificial neuron to values between two asymptotes. The sigmoid function is the most commonly used activation function. It is a continuous function that varies gradually between two asymptotic values typically 0 and 1 or -1 and +1.

4.2 Neural Fitting Tool

In fitting problems, you want a neural network to map between a data set of numeric inputs and a set of numeric targets. Examples of this type of problem include estimating house prices from such input variables as tax rate, pupil/teacher ratio in local schools and crime rate (house_dataset); estimating engine emission levels based on measurements of fuel consumption and speed (engine_dataset); or predicting a patient's bodyfat level based on body measurements (bodyfat_dataset).

The Neural Network Fitting Tool will help you select data, create and train a network, and evaluate its performance using mean square error and regression analysis.

A two-layer feed-forward network with sigmoid hidden neurons and linear output neurons (newfit), can fit multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer.

The network will be trained with Levenberg-Marquardt back propagation algorithm. (trainlm), unless there is not enough memory, in which case scaled conjugate gradient back propagation (trainscg) will be used. During past couple of years, the Levenberg-Marquardt (LM), a second order optimization technique is extensively employed in evapotranspiration modeling using neural networks.

In the present study, the meteorological data for the Rabi season are collected and correlated with yield of wheat for the period 1981- 1999 and validated for the 2000-2006 periods. The whole data is divided into 70% and 30% Datasets. The first part, i.e. 70% of the Dataset, is further divided into 70% for Training, 15% for Validation and 15% for Testing. For these datasets, correlations of input and output are observed using neural fitting tool. Then, the model is re-trained until the best coefficient of correlation is obtained and this corresponding model is considered as the best model and this is further used to validate the remaining 30% of the Dataset.

This whole procedure is repeated for three different Alternatives. In Alternative 1, only maximum and minimum temperatures are correlated with yield data. In Alternative 2, maximum and minimum temperature and relative humidity are correlated with yield data. In Alternative 3, Maximum and minimum temperatures, relative humidity and sunshine hours are correlated with yield.

4.3 Goodness of Fit Test Parameters Coefficient of Correlation, R

Measure of the "goodness of fit" is the coefficient of correlation, R to explain the meaning of this measure, one has to define the standard deviation, which quantifies

the spread of the data around the mean:

$$s_t = \sum_{i=1}^n (\bar{o} - o_i)^2$$

Where s_t is the standard deviation, o_i is the observed data points and \bar{o} is the average of predicted data points and the average the observed data points given by,

$$\bar{o} = \frac{1}{n} \sum_{i=1}^n o_i$$

The quantity s_t considers the spread around a constant line (the mean) as opposed to the spread around the regression model. This is the uncertainty of the dependent variable prior to regression. One also defines the deviation from the fitting curve as

$$s_r = \sum_{i=1}^n (o_i - p_i)^2$$

Where s_r is the deviation from the fitting curve, p_i is the predicted data points.

$$R = \sqrt{\frac{s_t - s_r}{s_t}}$$

where R is defined as the coefficient of correlation. As the regression model starts improving describing the data, the correlation coefficient approaches unity. For a perfect fit, the standard error of the estimate will approach $s_r = 0$ and the correlation coefficient will approach $R = 1$.

Note the similarity of this expression to the standard error of the estimate; this quantity likewise measures the spread of the points around the fitting function. Thus, the improvement (or error reduction) due to describing the data in terms of a regression model can be quantified by subtracting the two quantities. Because the magnitude of the quantity is dependent on the scale of the data, this difference is normalized to yield .

V. RESULTS AND ANALYSIS

The predefined correlation coefficient, R is found for each stage (training, validation, testing) for each model and these values also were found for over all data in addition to an additional results for a randomly selected data for additional testing .Values of Correlation coefficient, R are plotted for (training, validation, testing) for each model and are given in Fig. 1, 2 and 3 for Alternatives 1, 2 and 3 respectively.

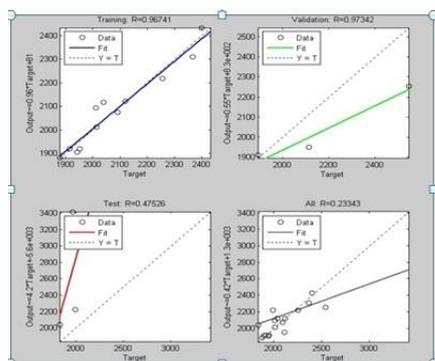


Fig.1 Correlation coefficient, R for Alternative 1

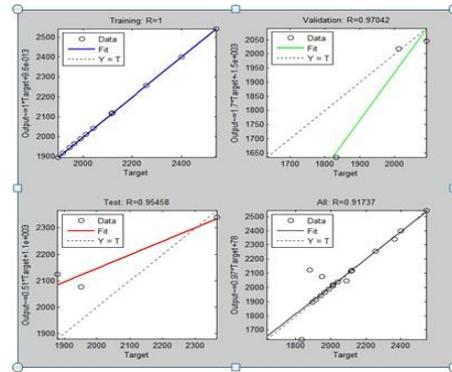


Fig.2 Correlation coefficient, R for Alternative 2

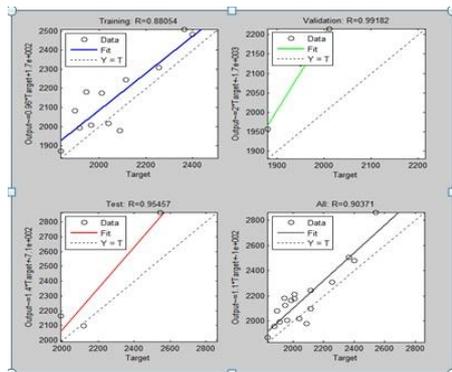


Fig.3 Correlation coefficient, R for Alternative 3

The Coefficient of correlation, R for 70 % and 30 % data set are given in Table 1.

Table 1 Coefficient of correlation, R for 70 % and 30 % data set

Alternative	Co- Efficient of Correlation, R			
	For 70 % of Dataset		For 30 % Dataset	
	Training	Validation	Testing	Validation
1	0.96	0.97	0.47	0.39
2	1.00	0.97	0.95	0.62
3	0.88	0.99	0.95	0.66

Looking to the Table 1 and Figs. 1,2 and 3, as per Alternative 1, i.e. considering correlation between maximum and minimum temperatures and Yield, the co-efficient of correlation for training and validation are better compared to testing for 70%, whereas in 30% dataset, R comes out to be 0.39, which is comparatively low.

As per Alternative 2, i.e. considering correlation between maximum and minimum temperatures & Relative Humidity and Yield, the co-efficient of correlation for training and validation are better compared to testing for 70%, whereas in 30% dataset, R comes out to be 0.062, which is comparatively good. As per alternative 3, i.e. Considering correlation between maximum and minimum temperatures & relative Humidity & Sunshine Hour and Yield, the co-efficient of correlation for training, i.e. 0.88 for validation, i.e. 0.99 an for testing for 70%, i.e. 0.95, whereas in 30% dataset, R comes out to be 0.66, which is good. Here, one can observe that sunshine hours are not

significantly affecting the yield of a crop compared to the effect of Maximum and Minimum Temperatures and relative humidity on the same yield of a crop.

Here, it is very much evident that considering maximum and minimum temperature and relative humidity variables, the correlation is achieved as the best, which reveals that yield of a crop is very much depended on maximum and minimum temperatures & relative humidity. Inclusion of Sunshine Hours does not play significant role.

VI. CONCLUSIONS

The correlation between maximum and minimum temperatures & relative Humidity and Yield, the co-efficient of correlation for training, i.e. 1.00 validation, i.e. 0.97 and testing, i.e. 0.95 are best for 70%, whereas in 30% dataset, R comes out to be 0.62, which is good.

The inclusion of sunshine hours with maximum and minimum temperatures and relative humidity does not show better correlation with yield of a crop compared to the consideration of only maximum and minimum temperatures and relative humidity.

It can be evidently concluded from the study that considering three variables, the correlation is achieved as

the best. This reveals that yield of a crop is very much depended on maximum and minimum temperatures & relative humidity.

The use of Neural Network fitting is quite helpful in studying the predominance of any variables on one another.

REFERENCES

- [1] Khashei-Siuki,A., Kouchakzadeh,M. and Ghahraman,B. "Predicting Dryland Wheat Yield from Meteorological Data Using Expert System", Khorasan Province, Iran, *J. Agr. Sci. Tech. (2011) Vol. 13: 627-640*
- [2] Imran Maqsood, Muhammad Riaz Khan and Ajith Abraham (2004). "An ensemble of neural networks for weather forecasting", *Neural Comput & Applic (2004) 13: 112-122*.
- [3] Parekh, F.P. and Suryanarayana, T.M.V. (2012). "Climatic Variability and Crop Weather Model of Wheat Crop for Vallabh Vidyanagar", *The Indian Journal of Technical Education, (Special Issue for NCEVT'12), April 2012, 325-330*.

An analysis on prioritizing tourism capacities of rural regions of Izeh city using TOPSIS model, Iran

Dr. Saeed Maleki¹, Davod Hatami², Mahdi Jadidoleslam³

¹Assistant Professor, Department of Geography and Urban Planning, Faculty of Earth Sciences and GIS, Shahid Chamran University of Ahvaz, Golestan St, Ahvaz, Iran.

²MA student of Sistan and Bluchestan University, Iran.

³Islamic Azad University of zahedan, Faculty of Science, Department of Geology, Iran.

Abstract: Rural tourism is a new tendency toward the industry of tourism. That has been brought forth since 50 s. nowadays, rural tourism as a strategy can undertake an effective role in developing and diversifying regional economy by proper scientific panning and identifying the potential and limitations of rural tourism. With this perspective, the obvious role of tourism in "Izeh" rural regions can be a movement toward regional balance. However, what is important in this social and economic movement procedure is the investment in places with touristic advantages. But it will not be achieved unless a proper selection is made among touristic regions so that the ground for progress of the whole area is prepared .in this procedure, use of techniques and models of prioritization is necessary. Present article, by laying the question, which one of studied rural regions is advantageous on ranking the potential priorities of tourism, tries to find a scientific answer, using analytical –descriptive method and TOPSIS model. The results show that among four sample villages of tourism (koolfarah, Shivand, Soosan and Noorabad villages), koolfarah with 0.895 coefficient possesses the highest rank and Noorabad with 0.427 coefficient possesses the last rank in investment.

Keywords: Tourism, Rural tourism, "Izeh" city, Prioritization, TOPSIS technique.

I. Introduction

Tourism is a franis term, originated from "tour" meaning "rotational movement", "the act of passing", "passing the environment", "to travel" and "walking". According to "pearlarous", tourism is to travel for enjoyment and tourism is to travel for enjoyment and tourist is a person who travels for enjoyment and satisfaction (Mahallati, 2001:2) for the first time in 1811, the term of tourism appeared sporting magazine that implied the travel in order to visit historic places and natural landscapes (Mahdavi, 2004:11), since 1970 s, studies on industry of tourism in their modern concept have gained a certain status among other academic sciences. Nowadays tourism is one of the most promising activities, that is considered as a transition to development (Zarabi, Islamiparikhani, 2011:2).21th century is the century of utilizing valuable trade oppurtunities in service sector, especially the tourism (Bahrami, 2000:2) nowadays, development of tourism in all domains whether at regional and national or at international levels is under consideration of public planers and private institutes. Many countries have realized that for improving their economic situation, they should take an initiative and try to find new ways. Awareness or the face that tourism is a remarkable resource of currency in come for economy has caused that tourism finds a wide concept in different economic, social and cultural dimensions and it is considered as an industry. Many countries know this dynamic industry as main resource of revenue, occupation, private sector growth and infrastructural development it is worthy to note that tourism is always considered as an important factor for economic development (Tayyebi, Babaki, Jabbari, 2007:2). Consequently, development of tourism should be investigated as a strategy from different economic, social, ecologic and foundational aspects. During recent 60 years tourism has grown continually and become one of economic sectors with a high growing rate in the world. During 1950 to 2005 international tourism increased by a rate of 6.5% from 25 million tourists to 806 million. Based in predications made by world tourism organization of united nations (NNWTO) the number of tourists of world will reaches 1.5 billion by 2020 (2006).

II. Theoretical bases

Tourism involves a set of activities tourists do in a period due to professional or personal reasons out of their working and living places (Smid, s and Zwart, 2002:23) WTO knows the tourism as a set of works a personal does in travel out of their residential place and it is not considered, more than a year, and its aim is enjoyment, entertainment, rest, sport and so kind activities. Industry of tourism is an integration of different activities which are carried out as a chain in order to serve the tourist. So tourism involves all phenomena and relations resulted from tourist's interaction with providers and sellers of touristic products, governments, and communities which are the hosts in the processes of attraction and welcoming the tourists (Mcintosh, 1995:9). In recent years due to different reasons of development (economic , social and ecologic) following the acceptance of this introducing, tourism has been considered as a tool for increasing the development of regions , economically and socially deprived (Pilar, Teresa, 2005:951), it has been increasing at international level regarding the number of tourists and revenue of currency (Pigozi B. and sambrook,2005:45). According to world tourism organizations statistical report, four occupational opportunities are created per a tourist who inters the region (WTO, 1999). The number of tourists is growing rapidly. Based on WTO reports, France with 75 million tourists and America with anemone about 74.5 billion \$ are standing at top position in Tourism. Among Mediterranean countries, turkey has been record bearer in recent four years (Goosling, 2004:45). Tourism is one of global streams which properly indicate integration of economic social and

cultural affairs (Holjeva, 2003:132). Tourism is a tendency to use the space in free time with different objectives and motives.

Tourism is a generality, involving a stream of capital, human and culture and interaction of them that leaves different effects in geographical spaces. Creation of wealth and occupation for local residents in delivering space for tourists use (Briedenhann, 2004:7). Emphasis on economic terms in geographic spaces for development and welfare of local residents substrates the tourism. To make the tourism a global issue necessitates macroeconomic policies at global level (Sugiyatro, 2003:683-684). The term of rural tourism has progressed since 1950s. And since 1960s and 1970s, economy of rural tourism for agricultures and local communities has been more considered. Until now, with presenting different strategies, the planers and experts have taken wide measures in attracting tourists and development touristic activities (Bahrami, 2010:3). World tourism organization as international foundation for tourism has realized the importance of rural tourism and held many seminars in Europe on analyzing the present and future condition recently. Rural tourism in fact includes various types such as green tourism, agricultural tourism, advantural tourism sport tourism and cultural tourism. This different typology causes difficulty in presenting a certain introduction of rural tourism (Rosa, Pabi, Andlibia, 2007:951). However, by accepting the introduction issued by WTO, it can be said that rural tourism beside activities like agriculture has had a proper effect and provided the villagers with important secondary income (Wilson et al, 2001:132). Generally, rural tourism is important from two aspects, first one as a wide global activity and the second as an emphasis on development of local and regional policy (Ahmadi, 2001:2260).

In 1986 European Union committee introduced rural tourism as follows: rural tourism does not include only agricultural tourism but it includes all touristic activities in rural regions (Mahdavi, 2004:18). Based on definition of Dot, rural tourism is everything that attracts the tourists toward surrounding areas of mother city (Knowd, 2001:90)

Rural tourism is referred to all services and activities which are performed by agricultures, people and government for tourist's entertainment and relaxation and also tourist's attraction to rural areas.

It can involve agricultural tourism, farm tourism natural and cultural tourism (Mahdavi, 2004:18).

With respect to what was said, it can be stated that rural tourism is a new term in rural development literature that like development, has various dimentions and effects. Therefore, development of rural tourism is appropriate to a share it reforming the villages socially and economically (Munshizadeh and Nasiri, 2001:36). With respect to mentioned definitions, it can be said that rural tourism, by providing new oppurtunities for many villages, not only gives new life to villages but also develops these areas and survives these residential (Eftekhari, Mahdavi, 2005:8).

In the last decade of 20th century, many economic- social planers in European countries defined the tourism as a confidential method with rightful perspective for development of villages especially the most deprived of them (Efteghari, 2002:32).

Studies in France, Austria, Swiss, England, inland, Toyland and Japan show that rural tourism grows rapidly in rural economy and it is complement to agricultural activities. The nature of tourism industry is creation of job and revenue, diversification of economy, social participation and use of local resources. Since a good deal of problems, depravedness and development is related to the lake of this industry, rural tourism, by solving those problems and constraints can lead to rural development (Papoliyazdi and Safaie, 2006:32). So development of tourism in rural areas can have an important role in diversification of economy of rural communities and can provide the ground for sustainable development on one hand it is a tool for motivating the growth of national economy through coping on envelopment perceptions and improving local people life standard.

On other hand, non-planned development causes social-environmental damages in rural residential. Therefore, with respect to environmental conditions, eighter a certain kind or combination of strategies should be utilized for development of rural regions.

III. Methodology

Success and sustainability of tourism depends on proper selection of touristic regions, considering capabilities and potential through preparing effective indexes in prioritization. According, present article tries to prioritize rural touristic region, using TOPSIS method, in order to select correctly touristic sample regions, improvised investment and realization of the aims of rural development of studied region. This research is a field study with descriptive-analytical method. To prioritize touristic capacities of rural regions of "Izeh", multi-variable mathematical planning and TOPSIS model were used. Multi-variable planning includes a wide range of mathematical techniques, different methods of which are used based on the aims of study. It involves two main groups of multi-aims planning and multi-variable planning. With respect to different indices related to prioritization of touristic capacities of rural regions, TOPSIS method was used in this research which is a sub-set of MADM.

1.1. Geographical location Izeh city

Izeh, whit on area about 3863 km² on eastern north of Khuzestan province is located between latitudes of 49°-45', 50° eastern and latitudes of 31°,45'-32° northern, mean altitude of plain is 830 m° Izeh has the coldest climate in in the province. It possesses different touristic potential, including 48 national recorded places, most of which are located in rural regions.



Figure I: the region of study (“Izeh”)

1.2. TOPSIS model

TOPSIS as a multi-variable decision making method is a simple but effective in prioritization. This method, for the first time in 1981 was set forth by chin and Howang, referring to Howang and yons book (Seraphim opriciuic, 2006: 665). TOPSIS technique belongs to compensative models (These models are important in exchange between indices) and adaptive sub- group - it is a preferred option which is the closest option to ideal solution. TOPSIS algorithm is a powerful compensative multi-indices decision making for prioritizing the options through simulation to ideal answer. In This method selected option should have the closest distance to ideal answer and the most remote distance to the most ineffective answer (SHayan, 2000:3).

The most important advantage of this method is that subjective and objective criteria and indices can be used simultaneously (Jhon, 1988:2).however, it is suggested that TOPSIS method should be used when the number of indices and available information is limited (Naumann, 1998:8). In utilizing this method, the following steps are implemented (Asgharpoor, 2008: 270-213).

IV. Discussions

Step1) formation of data matrix or decision matrix based on alternative n and index k.

$$A_{ij} = \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

W	40%	25%	15%	20%
The region of Tourism	historical cultural	geopolitics forms	number of natural	welfare facilities
soosan	60	6	4	50
village- shivand	70	5	3	60
village- noorabad	75	3	2	70
village- koolfarah	55	4	2	60

step2) in step 2, decision or data matrixes are normalized, and then standard matrix is formed. The formula for normalization is as follows:

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{k=1}^n a_{kj}^2}}$$

0.4052	0.5962	0.6469	0.4581
0.5267	0.4472	0.5391	0.5345
0.5673	0.2981	0.3234	0.5727
0.4862	0.5962	0.4313	0.4199

step3) Based on expertise approaches, theoretic facts and other methods as well as the importance of each criterion, the weight of each index is gained.

$$V_{ij} = \begin{pmatrix} w_1 r_{11} & w_2 r_{12} & \dots & w_n r_{1n} \\ w_1 r_{21} & w_2 r_{22} & \dots & w_n r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_1 r_{m1} & w_2 r_{m2} & \dots & w_n r_{mn} \end{pmatrix}$$

$w_1 = 20\% \quad w_2 = 15\% \quad w_3 = 25\% \quad w_4 = 40\%$

0.4052	0.5962	0.6469	0.
0.5267	0.4472	0.5391	0.
0.5673	0.2981	0.3234	0.
0.4862	0.5962	0.4313	0.

step4) determination of distance of I alternative from ideal alternative, saying determining the highest performance of each index, which is shown by (A^+) , and it is given by:

$$A^* = \left\{ (\max_i v_{ij} | j \in J), (\min_i v_{ij} | j \in J') \right\} \rightarrow A^* = \{v_1^*, v_2^*, \dots, v_n^*\}$$

$A + \{./1134 \quad -./0894 \quad -./1617 \quad -./2290\}$

Step5) determination of the lowest performance .by the following equation:

$$A^- = \left\{ (\min_i v_{ij} | j \in J), (\max_i v_{ij} | j \in J') \right\} \rightarrow A^- = \{v_1^-, v_2^-, \dots, v_n^-\}$$

$A - \{./0810 \quad -./0447 \quad -./0808 \quad -./1679\}$

Step6) determination of distance criterion for maximum and minimum alternative: calculation formula is as follows:

$$S_i = \sqrt{\sum (v_{ij} - v_{jt})^2}$$

V. Ideal alternative

$$./0924 S_i^+ = ./0830 = ./0390 S_i^3 = ./0561 S_i^2 = S_i^1$$

$$./0936 S_i^2 = ./0781 S_i^3 = ./0691 S_i^4 = ./714 = S_i^1$$

Step7) In this step, the equal coefficient is gained from dividing the minimum alternative in to (minimum alternative+ maximum alternative). In other words, relative closeness A: in ratio to A^* is calculated. Its equation is as follows:

$$C_i^* = \frac{S_i^-}{S_i^- + S_i^*}$$

Soosan village	$= \frac{./0936}{./0936 + ./0561} = \frac{./0936}{0/1497} = 0/625$
0/666shivand village	$= \frac{./0781}{./0781 + ./0390} = \frac{./0781}{0/1117} =$
Noorabad village	$= \frac{./0691}{./0691 + ./0924} = \frac{./0691}{0/1615} = 0/427$
Koolfarah village	$= \frac{./714}{./714 + ./0830} = \frac{./714}{0/797} = 0/895$

Estimated indices of four touristic places in Izeh, using TOPSIS model show that koolfarah village with 0.427 score is ranked top which is followed by Noorabad with 0.666 and shivand with 0.895 in attracting tourists based on TOPSIS model.

Hist of Touristic	Priority
koolfarah Village	1

shivand Village	2
nooraload Village	3
negarandegan Village	4

VI. Conclusion

The industry of tourism, as ultra-sector industry involves different activities. It has an interaction with all economic, social and cultural sectors directly or indirectly. Consequently, one of the fundamental factors in creating evolution and alteration in the county's tourism sector is comprehensive study of positive and negative effects and consequents and regional insight into its development.

It means that tourism should be planned in the mould of systematic procedure for development and flourish merlot. Rural tourism in development rural local communities in "Izehh" results in independence and empowerment of region economy, flourishing of handicrafts, increase of life quality, and solution of joblessness problem. Analysis of tourism effects and consequents in rural regions of "Izehh" shows that positive aspects related to tourism have shown their effects and have been followed by desirable consequents including the role of huckster's business in increasing villagers in come, flour is hment of gardening, improvement of cultural , reduction of villagers migration to cities , flourish meant of handicrafts market like shoe weaving , consider elation to traditions and customs in holding nations and religious festivals . At mean time, the limitations before the development of this industry are: lake of province authorities attention to revenue creating role of this in destroy, in sufficient bud yet in developmental part, weakness of in frasrctures of residential services and communicational routs nut work and so on.

One of the methods of multi- criterion decision making is mentioned in this article multi-criterion decision-making method addresses the issues that transit from various and even conflicting criteria. In fact, decision – making is selection process of being single optional among existing options. By in trod using the weights, implementation of TOPSIS a logarithm is dealt with, while the logarithm prioritizes dif ferment regions with respect to cts' work base. The results of study show that TOPSISa logarithm is a powerful decision – making techniqeal for prioritizing of options through simulating to ideal answer.

It is obvious that regarding developmental in dices (economic, social and ecological) and temporal cap acuity measurement, investment in a touristic region results in desirable consequents and provides the ground for total development, but it is not possible an less scientific methods are used in prioritization so that a proper selection among touristic sequins is made . As previously mentioned that concrete identification of regions potentials and capacities and their prioritization are necessary for proper selection of sample regions in order to correct incest and to realize rural development. Accordingly, four villages with touristic attraction in "Izeh" were selected and prioritization method of TOPSIS was carried out. Based on arising question and precise planning of touristic development in this region, four induces of cultural social attractions, geotouristic forms, number of natural landscapes and well-fare utilities were used as main criterion of measuring prioritization of rural touristic capacities, in "Izeh". Based on TOPSIS method, prioritization of mentioned villages is defined as follows. Koofarah village was identified as first priority while Noorabad village was known as last priority. By utilizing TOPSIS method in villages with touristic attractions, it was known that koofarah village has the highest chance in attracting tourist and investment. It is necessary that authorities should pay attention to investment and creation of proper infrastructures with respect to insufficient budget assigned to this field, and by sufficient and timely investment, should provide the ground for attracting tourists. the results of research show that "Izeh" com host a lot of national and international tourists , due to existence of touristic attractions and cultural values, by investing in which sustainable development is possible, under in flounce of tourist attractions in "Izeh" city.

In other words, if needed infra- structures for using these attractions are provided, occupation and income achievement for residents will be possible.

References

- [1] Abby Liu, Geoffrey Wall (2006). Planning tourism employment: a developing country perspective, Tourism Management, Vol 40, No 27.
- [2] Asghar poor, Mohammad Javad, (1998), multi-criteria decision making, Tehran, Tehran University Press.
- [3] Briedenhann, Jenny and Eugenia Wickens , (2004). Tourism Routes as a tool for the Economic Development of Rural Areas – Vibrant Hope or Impossible Dream ? Tourism Management, Vol 25 .
- [4] Bahrami, Rahmatullah, (2010), review the capabilities and bottlenecks in rural tourism development in Kurdistan, Proceedings of the Fourth International Congress of Geographers of the Islamic world.
- [5] Cultural Heritage and Tourism Base ize, www.mpr_khuz.ir
- [6] Goossling, S C (2004) . Global environmental Consequences of tourism . global environmental –change . journal .vol12 : 45 – 67 .
- [7] Gemma canoves, Montserrat villarino, Gerda k. Priestley, Asuncion, Blanco (2004). Rural Tourism in Spain: an analysis of recent evaluation, Geoforum, No 35.
- [8] Holjeva, Ivanka (2003). A Vision of Tourism and the Hotel Industry in the 21 Century , Hospitality managemwnt. VOI 22 .
- [9] Hekmatnia, H., Mousavi, Mir Najaf, (2001), the use of models in geography with an emphasis on urban and regional planning, modern science publications, Yazd, 150.

- [10] Hekmatnia, Hassan and Mousavi, Myrnajaf, (2001), the use of models in geography with an emphasis on urban and regional planning, modern science publications, Yazd.
- [11] Isabel Pilar Albaladejo Pina, María Teresa D'az Delfa (2005). rural tourism demand by type of accommodation, Tourism Management, Vol 35, No 26.
- [12] Jenny Briedenhann, Eugenia Wickens (2004). Tourism routes as a tool for the economic development of rural areas vibrant hope or impossible dream, Tourism Management, Vol 32, No 25.
- [13] John F. Affisco (1988). An empirical investigation of integrated spatial-proximity MCDM-Behavioral problem solving technology group decision models, Developments in Business Simulation & Experiential Exercises, Vol 15, Hofstra University.
- [14] John, F. Affisco ., An empirical investigation of integrated Spatial – Proximity MCDM – Behavioral Problem Solving technology group decision models, Developments in Business Simulation & Experiential Exercises, Vol. 15, Hofstra University, 1988 .
- [15] Mahallati, Salah al-Din, (1991), Introduction to Tourism, Beheshti of University.
- [16] McIntosh, Robert. W, Goeldner, Charles and Ritchie, R. Brent. (1995). tourism, Principles, Practices, Philosophies, United States of America: John Wiley & Sons, INC.
- [17] Monshizadeh, Rahmatullah, Nasiri, F., (1380), rural tourism, publishing secretary
- [18] Naumann. Fles., "Data fusion and data quality" Institut für Informatik, Humboldt – Universität ZU Berlin, 1998 .
- [19] Naumann Flex (1998). "Data fusion and data quality" Institut für Informatik, Humboldt - Universität zu Berlin.
- [20] Papy Yazdi, Mohammad Hossein, Saghaei, M., (2006), tourism (nature and concepts), the Press, 85.
- [21] Qryshvndy, Ismail, Qryshvndy, Hamid and Hatami, Davod (2010), examining the potential of tourism in city, Khuzestan province in the North East, Fourth National Conference on Geological Advertising Khorasan.
- [22] Rukn al-Din, Eftekhari, A., Mahdavi, D., (2005), rural tourism development strategies using SWOT Lavasan small villages.
- [23] Rukn al-Din, Honorable, Ismail, (2002), the role of rural tourism in rural development, Journal of Madras, (Humanities), Volume VI, Issue II, University of Tehran.
- [24] Rahimi, Hosain, (1992). Rural tourism and its territory, economic and political magazine, No. 185 186, 226.
- [25] Sharif-Zadeh, Abolghasem., Humayun Muradnejad, (2002), sustainable development of rural tourism, Journal of Social, Economic Jihad, No. 250 251, 52.
- [26] Statistical Center of Iran, (2006), General Population and Housing Census.
- [27] Shayan, Ali, (2010). I use multiple criteria decision making techniques in selecting the appropriate strategy for implementing IT projects, Iran's Industrial Management, 1385.
- [28] Sugiyarto, Cuntur and Adam Blak, (2003). Tourism and Globalization, Annals of Tourism Research, Vol 30 .
- [29] Serafim, O., Gwo-Hshiung, T., Compromise Solution by MCDM Methods: A Comparative Analysis of VIKOR and TOPSIS ; European Journal of Operational Research, 156, 445 -455, 2004 .
- [30] Zarrabi, AA, Islami Prikhany, Sdyf, (2011), measuring the impact of economic, social, cultural and ecological tourism development, (case study city Meshkinshahr), Human Geography Research, No. 75, 3752.
- [31] Tayebi, Kamil, Babaki, Ruhollah and Jabbari, Amir, (2007), examining the relationship between tourism development and economic growth in Iran (13,381,383), Humanities and Social Studies Special Economic Letter, Year Seven, Number Twenty-Six, fall 1386.
- [32] The United Nations World Tourism Organization (UNWTO) (2006) Tourism Highlights, Facts & Figures section at www.unwto.org .
- [33] Ziari, Keramatallah, (1383), schools, theories and models and Regional Planning Program, University of Yazd, Yazd printing.

Identification of QR Code based on Pattern Recognition with Mobile Phones

Shanthi Kumaraguru, Prof. Dr. D. S. Bormane

JPMS's Rajarshi Shau college of Engineering.

Department of Electronics and Telecommunication Engineering, Thathawade, Pune-44, India

Abstract: Quick Response Code has been widely used in the automatic identification field. This paper proposes a implementation of real-time Quick Response Code recognition using mobile, which is an efficient technology used for data transferring. An image processing system based on mobile is described to be able to binarize, locate, segment, and decode the QR Code.

Keywords: Quick Response Code, Binarization, recognition, mobile- Phone

I. INTRODUCTION

Bar code is fast, easy accurate and automatic data Collection method. Bar code enables products to tracked efficiently and accurately at speed not possible Using manual data entry System. The bar code reader only be used to recognize the bar code, and the price of two dimensional bar code reader is expensive.

Now mobile phones can implement much new kind of new applications such as taking photos, and movies shooting by using embedded camera device. So and interesting approach is capturing bar-code with their cameras and decoding them with software running on the phone. Recently the mobile industry begins to pay more attention to bar code application in m-commerce because 2D barcodes not only provide a simple and expensive method to present diverse commerce data, but also improve mobile user experience by reducing their input.

Previous research work has shows that recognition of 2D barcode is researching hotspot and very difficult in various conditions. These condition include highlight spots, skew projections low contradistinction, non-homogeneous lighting, and various mixed conditions.

Ohbuchi et al presented an algorithm capable of the real-time recognition of barcodes on a mobile phone. The most important drawback of the method is the fact that it has been hand tailored for one certain hardware device. It relies on access to a powerful but also very specific hardware element, which is not accessible by normal application developers: the built-in signal processor of a device. Besides that, the algorithm proposed by Ohbuchi relies on two specific conditions: Prior to the code's decoding, the barcode's position is detected using a spiral scanning algorithm that runs on the device's signal processor.

This algorithm makes the assumption that the point in the middle of the screen is located in the code. This assumption is not realistic in usually environment. Ying et al presented an algorithm capable of recognizing the PDF 417 barcodes in real time with a mobile phone. This method includes three steps. The first step detects the code region using the Otsu algorithm and the Least Square Method.

The Second method searches for the cut-off rules with a scanning approach. In the third step symbol characters are segmented from the original image. So the successful linearization of the code areas during the first step is the most key step in his algorithm. This is mean the recognition result will rely on the effect of the Otsu method. But the Otsu method could be failed in complex lighting conditions, including highlight spots, low contradistinction, non-homogeneous lighting, and various mixed conditions. so the recognition rate of the algorithm proposed by Ying is not high especially in complex lighting conditions. Additionally, many constraints Are added in this method, so the application scope is decreased.

Sun et al. Introduce an algorithm to analyze and correct the distorted image of QR code. The algorithm consists gray-scale image transformation, Canny edge detection, external contours finding, inverse perspective transformation and cell grids generating. In this method, the recognition time will be cost more than Ying's method because there is no binarization operation; gray image is processed in all steps of Sun's method. The Processing time of the black-white image is less than the processing time of gray image.

II. THE STRUCTURE OF QR CODE

Quick Response Code is developed in Japan by Denso Corporation in 1994, and later recognized as a standard. QR code has been approved as an AIM Standard, a JIS standard and an ISO standard. So QR Code is used in variety of applications, such as manufacturing, logistics, and sales applications. There are 40 versions in QR Code, four levels of error correction, and the maximum symbol size can encode 7089 numeric data or 4296 alphanumeric data. The Reed Solomon is used in QR Code, and the highest level of error correction allows recovery of 30% of the symbol code words. In addition to, QR Code has a clear feature that a finder pattern in the upper left and right and lower left corners.

Each QR Code symbol consists of an encoding region and function patterns as shown in fig.1. Function patterns include finder, separator, timing patterns and alignment patterns. The Finder pattern located at the three corners of the symbol intended to assist in easy location of its position, size and inclination.

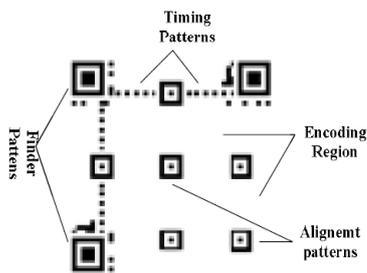


Fig.1.The structure of QR Code

III. OUR APPROACH

It is proposed to devise the recognition algorithm of QR Code which can be used in various lighting conditions. The algorithm consists of steps as gray scale image conversion, binarization, filter, orientation (finder patterns or timing patterns location), alignment patterns location, cell grids generating, error correction and decoding.

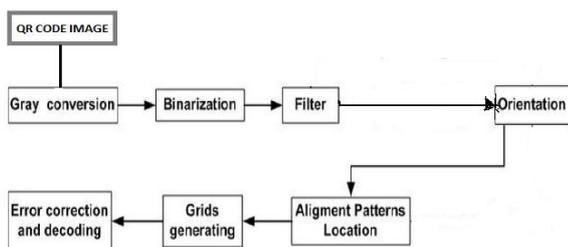


Fig.2. Flow chart of the proposed work

3.1 Gray Conversion

QR Code symbol is captured by mobile phone with camera, and images are captured in RGB 24bit format by most phones, but QR Code symbol is a set of dark and light pixels. It is needless to deal with color information and the gray image calculated quickly with little space, so gray conversion is needed to do firstly.

3.2 Binarization

Selection of a proper binarization method is critical to the performance of barcode recognition system. In binarizing an image, a simple and popular method is threshold. Among more than 20 threshold methods concluded that Otsu's method is the best, which chooses the threshold that minimizes within-group variance. But in our test, Otsu's algorithm is not satisfactory performance in uneven illumination and is not real-time implementation. Ohbuchi [3] propose a global threshold method. First he divided the center 60*60 areas of 320*240 images into nine blocks. Then calculate the gray histogram for each block, and sort the gray value. The middle value is chooses as the threshold of each block. Finally the smallest value of these middle values is the global threshold of this whole image. This method is existed two drawbacks: (1) the barcode symbol must be in the center of the captured image, otherwise the method will failed. (2) This method Result in excessive segmentation. Separate a part of barcode image as background and lead to decode failure.

Using a global thresholding method, if an image has variable lighting conditions, the resulting binary image will be very bad . In this case, a local thresholding method performs better. The main problems with a local threshold method are hard to set a right window size, eliminate the

block effect, and reduce the execution time. However, memory restrictions and embedded system requirements preclude the use of binarization algorithms that require a priori knowledge of the full image and large execution time, thus a number of well-known locally adaptive algorithms cannot be use. So it is difficult to binarize the bar code in various illuminations using one method.

An adaptive multi-level thresholding algorithm is purposed, which integrate the local threshold with global threshold. The algorithm can achieve higher recognition rate under the condition of lower illumination, contrast and uneven illumination. This method including following steps. Firstly, calculate the histogram of gray image. In order to decreasing the effect of noise, filter the histogram, and analysis the feature of histogram peak. If the filtered histogram is bimodal distribution, the lowest of trough or the middle value of flat trough will be used as the global threshold. In common lighting conditions, this global threshold is used to process image with the satisfying result. If the histogram of image is single peak histogram, and single peak area in lower gray area, it means that barcode image is in weak illumination; otherwise the barcode image is in strong illumination. We adopt iterative threshold method, which form the threshold with mean between old global threshold and center of dark area or light area. If histogram shows multipeak distribution, it means that the image is in the case of uneven lighting condition or complex background. The local threshold algorithm is used. The multi-level threshold method integrate global threshold and local threshold method it is meeted the real-time binarization in common lighting condition[1], and also satisfied the binarization in special illumination condition.

3.3 Filter

Standard opening and closing techniques are applied to the bitmap to remover noise. After the filter, edge detection is used in most recognition algorithm. But the QR code has its special features; three finder patterns will provide the quick orientation. So the step of edge detection is omitted in our approach as shown in figure this will increase the recognition rate.

3.4 Orientation

There are three identical position detection patterns located at three of the four corners of QR Code, as shown in figure 1. Three dark-light-dark squares are overlapped in very finder pattern, and the dark-light ratio is 1:1:3:1:1. There is the off chance that similar graphic existed in barcode symbol. So the approximately ration area should be quickly find. But when one of the finder patterns is partially dirty or damaged, we can use the timing patterns. The timing patterns provide the secondary information which can help us to locate the symbol, decide the rotation direction of symbol and the width of module . A, modular distance offset algorithm was used without rotating symbol. After located the finder patterns, the module width and height and the angle of rotation were known. We get the grid moving along the line which connects the central of three finder patterns. This method avoids the rotation and interpolation, increases the computing speed.

3.5 Alignment Patterns Location

Different from PDF 417 barcode, in order to correct the contorted QR code symbol, there are many alignment patterns in symbol. With the version increasing, the number of alignment patterns adding, as shown in fig 3. When version is 3, the number of alignment patterns is 1, but when version is 7, the number of alignment patterns is 6. Link the central point of the alignment patterns and three position detected patterns, the small sampling grid is formed. In small sampling grid, distortion is omitted. Therefore, located the central coordinates of alignment pattern is critical for recognition barcode.



Fig. 3 The number of alignment patterns

Using the known alignment patterns and detection patterns can estimate the central coordinate. In international standard of QR Code, a reference method that locates the central coordinates of alignment patterns is provided. But the precondition of this method is estimated the central coordinate just inside of the alignment patterns. A large of actual captured image are tested, the probability of estimated central value inside of the alignment patterns is only 70% . The other 30% image will be located failure using these methods. A new estimated method is provided which used eight directions of the estimated point. The method scan the nearly eight directions of the estimated point to find the central coordinate of alignment patterns which are satisfied condition. This method expands the scanning range, and 99.3% captured QR code image is successfully decoded.

3.6 Grid Generation

When the finder patterns and alignment patterns were located successfully, the segment is easy to do, and the grid is easy to generate. Then the corresponding pixels were getting to ready decoding.

3.7 Error Correction and Decoding

The error correction and decoding process is the last step of recognition barcode. It employs the Reed-Solomon error correction to enable accurate reads even when substantial parts of the code are distorted. Decoding is just the reverse of the encoding procedure and the decoding steps can reference to international standard of QR Code.

IV. RESULTS

A user can scan and decode QR Code using a mobile phone with a built-in camera and Decoder software. After decoding the QR Code symbol, the user is forwarded to product related web-sites or marketing campaigns. The application software developed in java is executed in the computer and when it is ready to process ,

executable Jar File is copied to the Mobile Phone which is with Symbian Operating System. Then the image on any surface like NewsPaper, on the Computer ,Magazines are captured and Decoded .

Sr.No	Noise Level in %	Decoding level in %
1	0	100
2	5	100
3	15	100
4	25	100
5	35	100
6	45	100
7	55	100
8	65	100
9	74	100
10	77	87
11	80	85
12	90	55
13	95	Decoding failed

Table – Result testing report for various noise levels:

V. CONCLUSION

With the Mobile Phone with camera device is getting more popular, recognition of barcode based on mobile phone is getting more important and practical, the method mentioned in this paper is a new high-speed, high-accuracy automatic recognition method for recognizing QR Code symbol in various illumination conditions. This method has no hardware specific requirements unlike ohbuchi's method and is able to run as a standard application on most of the mobile phones.

ACKNOWLEDGMENT

I would like to thank professor Dr. D. S. Bormane for guiding me in this work in RSCOE college, Thathawade, Pune.

REFERENCES

- [1] Yue Liu, Ju Yang, Mingjun Liu "Recognition of QR Code with Mobile Phones ," 978-1-4244-1734-6/08/2008 IEEE.
- [2] TasoFalas, Hossein Kashani, "Two-Dimensional Bar-code Decoding with Camera-Equipped Mobile Phones," Proceedings of the Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops(PerCom W'07), 19-23, March, 597-600, 2007.
- [3] Gao, Jerry Zeyu, et al, "Understanding 2D-BarCode Technology and Applications in M-Commerce", The 31st Annual International Computer Software and

Applications Conference(COMPSAC 2007). Volume 2, July 24-27, 49-56, 2007.

- [4] Eisaku Ohbuchi, Hiroshi Hanaizumi, and Lim Ah Hock, "Barcode Readers using the Camera Device in Mobile Phones," Proceedings of the 2004 International Conference on Cyberworlds(CW'04), 11, 260-265, 2004.
- [5] Robert Adelman, ETH Zurich, "Mobile Phones Based Interaction with Everyday Product – On the Go," The 2007 International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST 2007). 12-14, Sept., 63-69, 2007.
- [6] Ying-Hong Liang, et al, "Real Time Recognition of 2D Bar Codes in Complex Image Conditions," Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22, Aug, 1699-1704, 2007.
- [7] Yue Liu, Mingjun Liu, "Automation Recognition Algorithms of Quick Response Code Based on Embedded System," Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06). Volume 02, Nov., 783-788, 2006.
- [8] Aidong Sun, Yan Sun, "The QR-Code reorganization in illegible snapshots taken by mobile phones," Fifth International Conference on Computational Science and Applications. 26-29, Aug., 532-538, 2007.
- [9] ISO/IEC 18004:2000. Information technology- Automatic identification and data capture techniques-Bar code Symbology-QR Code, 2000.

Implementation of D-STATCOM for Improvement of Power Quality in Radial Distribution System

Kolli Nageswar Rao¹, C. Hari Krishna², Kiran Kumar Kuthadi³

ABSTRACT: D-STATCOM (Distribution Static Compensator) is a shunt device which is generally used to solve power quality problems in distribution systems. D-STATCOM is a shunt device used in correcting power factor, maintaining constant distribution voltage and mitigating harmonics in a distribution network. D-STATCOM is used for Grid Connected Power System, for Voltage Fluctuation, for Wind Power Smoothing and Hydrogen Generation etc. This paper D-STATCOM is used in Electrical Power System for Power Quality Improvement. Relevant solutions which applied nowadays to improve power quality of electric network according to the five aspects of power quality- harmonics, fluctuation and flick of voltage, Voltage deviation, unbalance of 3-phase voltage and current frequency deviation. The D-STATCOM injects a current into the system to mitigate the voltage sags. LCL Passive Filter was then added to D-STATCOM to improve harmonic distortion and low power factor. The simulations were performed using MATLAB SIMULINK version R2009a.

Key Words: D-STATCOM, Voltage Sags, Voltage Source Converter (VSC), LCL Passive Filter, Total harmonics Distortion (THD).

I. INTRODUCTION

In order to improve the survivability of a navy ship in battle condition, DSTATCOM or Distribution Static Compensator can be used, which reduces the impact of pulsed loads on the bus voltage and thus keeps the bus voltage at desired level. DSTATCOM is a voltage-source inverter (VSI) based shunt device generally used in distribution system to improve power quality. The main advantage of DSTATCOM is that, it has a very sophisticated power electronics based control which can efficiently regulate the current injection into the distribution bus. The second advantage is that, it has multifarious applications, e.g. a). Cancelling the effect of poor load power factor, b). Suppressing the effect of harmonic content in load currents, c). Regulating the voltage of distribution bus against sag/swell etc., compensating the reactive power requirement of the load and so on. The performance of the DSTATCOM is very much dependent on the DSTATCOM controller.

II. DISTRIBUTION STACTOM

i. Basic Principal of D-STACTOM: A DSTATCOM is a controlled reactive source, which includes a Voltage Source Converter (VSC) and a DC link capacitor connected in shunt, capable of generating and/or absorbing reactive power. The operating principles of a DSTATCOM are based on the exact equivalence of the conventional rotating synchronous compensator.

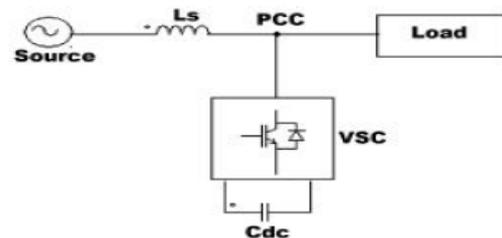


Fig. 1 Basic structure of D-STATCOM

The AC terminals of the VSC are connected to the Point of Common Coupling (PCC) through an inductance, which could be a filter inductance or the leakage inductance of the coupling transformer, as shown in Fig. 1. The DC side of the converter is connected to a DC capacitor, which carries the input ripple current of the converter and is the main reactive energy storage element. This capacitor could be charged by a battery source, or could be recharged by the converter itself. If the output voltage of the VSC is equal to the AC terminal voltage, no reactive power is delivered to the system. If the output voltage is greater than the AC terminal voltage, the DSTATCOM is in the capacitive mode of operation and vice versa. The quantity of reactive power flow is proportional to the difference in the two voltages. It is to be noted that voltage regulation at PCC and power factor correction cannot be achieved simultaneously. For a DSTATCOM used for voltage regulation at the PCC, the compensation should be such that the supply currents should lead the supply voltages, whereas, for power factor Correction, the supply current should be in phase with the supply voltages. The control strategies studied in this paper are applied with a view to studying the performance of a DSTATCOM for power factor correction and harmonic mitigation.

ii. Basic Configuration and operation of D-STATCOM

The D-STATCOM is a three-phase and shunt connected power electronics based device. It is connected near the load at the distribution systems. The major components of a D-STATCOM are shown in Fig. 2. It consists of a dc capacitor, three-phase inverter (IGBT, thyristor) module, ac filter, coupling transformer and a control strategy. The basic electronic block of the D-STATCOM is the voltage-sourced inverter that converts an input dc voltage into a three-phase output voltage at fundamental frequency. The D-STACOM employs an inverter to convert the DC link voltage V_{dc} on the capacitor to a voltage source of adjustable magnitude and phase. Therefore the D-STATCOM can be treated as a voltage-controlled source. The D-STATCOM can also be seen as a current-controlled source. Fig. 2 shows the inductance L and resistance R which represent the equivalent circuit elements of the step down transformer and the inverter will be the main component of the D-STATCOM.

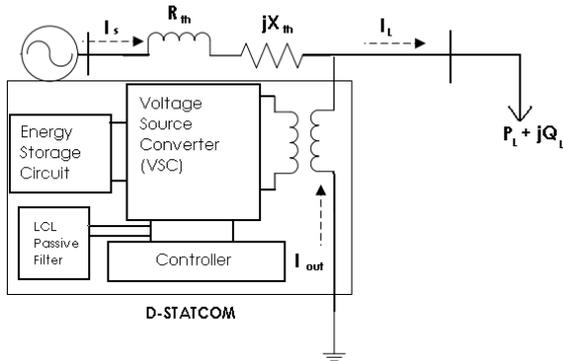


Fig. 2 Schematic diagram of a D-STATCOM

The voltage V_i is the effective output voltage of the D-STATCOM and δ is the power angle. The reactive power output of the D-STATCOM inductive or capacitive depending can be either on the operation mode of the D-STATCOM. The construction controller of the D-STATCOM is used to operate the inverter in such a way that the phase angle between the inverter voltage and the line voltage is dynamically adjusted so that the D-STATCOM generates or absorbs the desired VAR at the point of connection. The phase of the output voltage of the thyristor-based inverter, V_i , is controlled in the same way as the distribution system voltage, V_s .

iii. Compensation scheme of D-STATCOM

The D-STATCOM is a DC/AC switching power-converter composed of an air-cooled voltage source converter. Basically, the D-STATCOM is used to suppress voltage variations and control reactive power in phase with the system voltage. The D-STATCOM produces phase-synchronized output voltage, therefore, it can compensate for inductive and capacitive currents linearly and continuously. Active and reactive power trade between the power system and the D-STATCOM is accomplished by controlling the phase angle difference between the two voltages. If the output voltage of the D-STATCOM V_i is in phase with the bus terminal voltage V_t , and V_i is greater than V_t , the D-STATCOM provides reactive power to the system. If V_i is smaller than V_t , the D-STATCOM absorbs reactive power from the power system. Ideally, V_t and V_i have the same phase, but actually V_t and V_i have a little phase difference to compensate for the loss of transformer winding and inverter switching, so it absorbs some real power from system. Fig. 3 shows the D-STATCOM vector diagrams, which show the inverter output voltage V_i , system voltage V_t , reactive voltage V_L and line current I in correlation with the magnitude and phase α . Fig. 3(a) and Fig. 3(b) explain how V_i and V_t produce inductive or capacitive power by controlling the magnitude of the inverter output voltage V_i in phase with each other. Fig. 3(c) and Fig. 3(d) show that the D-STATCOM produces or absorbs real power with V_i and V_t having a phase difference $\pm\alpha$.

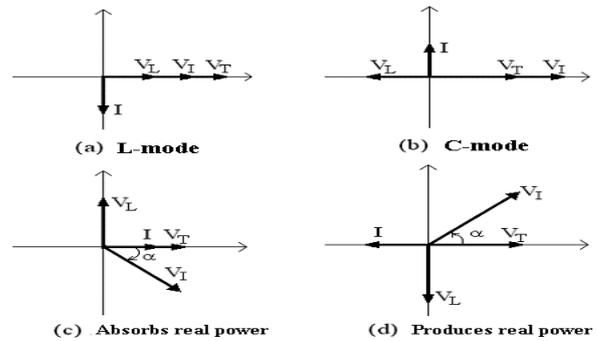


Fig. 3 Vector diagrams of D-STATCOM

Fig. 4 shows a radial type electric power distribution system feeding an unbalanced load. A DSTACOM is installed in parallel with the unbalance load for on-site load compensation. The reactive power output of the D-STATCOM in each phase, which is inductive or capacitive, can be independently controlled by the controller of the DSTACOM for real-time load compensation

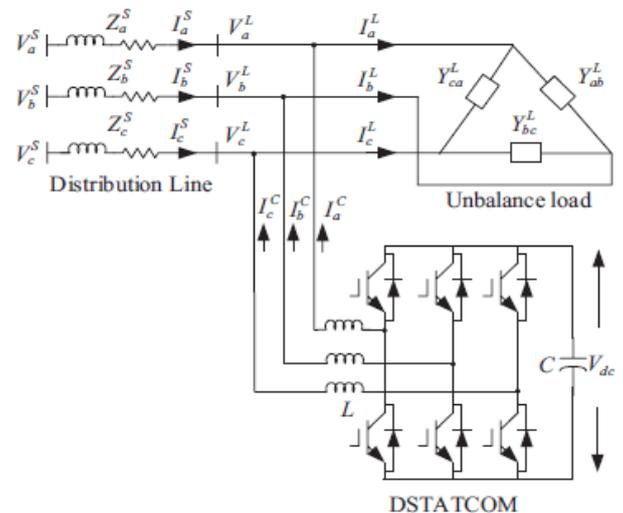


Fig. 4 A radial distribution system with an unbalance load and a D-STATCOM

III. Radial Distribution Test System

The test system comprises a 230kV, 50Hz transmission system, represented by a Thevenins equivalent, feeding into the primary side of a 3-wdg transformer connected in Y/Y/Y, 230/11/11 kV. A varying load is connected to the 11 kV, secondary side of the transformer. A two-level D-STATCOM is connected to the 11kV tertiary winding to provide instantaneous voltage support at the load point. A 750 μ F capacitor on the dc side provides the D-STATCOM energy storage capabilities. Breaker 1 is used to control the period of operation of the D-STATCOM and breaker 2 is used to control the connection of load 1 to the system.

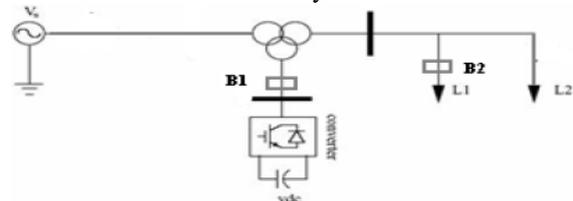


Fig. 5 Single Line Diagram of Test system

IV. Simulation Results & Discussion

4.1 Simulation Model for Test System without Insertion of D-STATCOM

To create distortion in the distribution system, different types of fault such as Three Phase to Ground (TPG), Double Line to Ground (DLG), Line to Line (LL), and Single Line to Ground (SLG) are injected.

Table 1 Results of voltage sags for different types of faults

R_f in Ω	SLG fault	DLG fault	TLG fault	LL fault
0.86	0.8679	0.7833	0.7515	0.8210
0.76	0.8485	0.7487	0.7106	0.7918
0.66	0.8259	0.7070	0.6600	0.7587

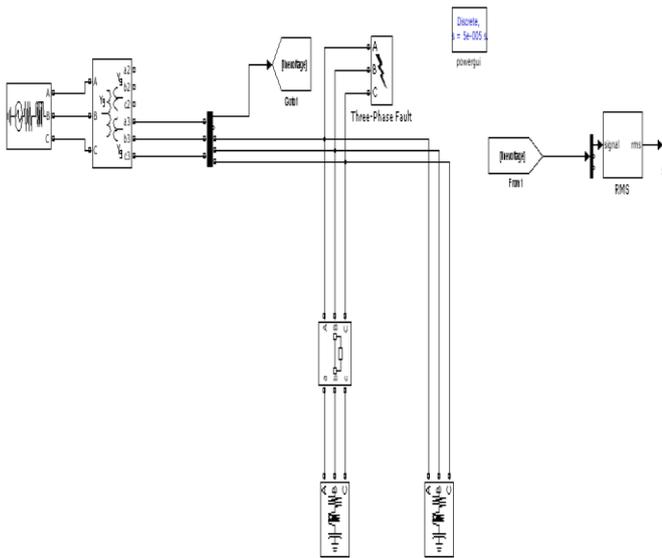


Fig. 6 Simulink Diagram for test system without D-STATCOM

Table 1 shows the overall results of voltage sags in p.u for different types of fault. From the table, it can be observed that when the value of fault resistance is increase, the voltage sags will also increased for different types of fault.

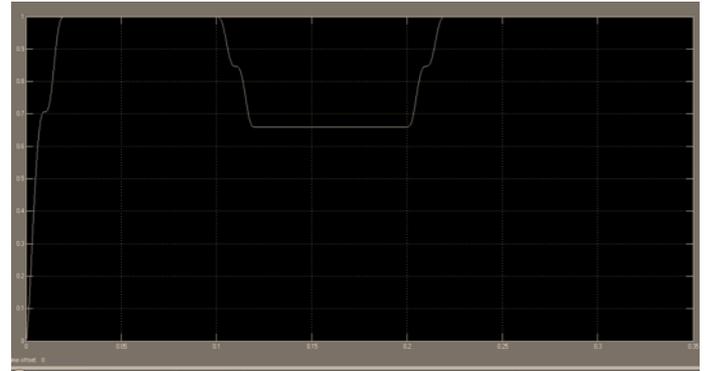


Fig. 8 Voltage at load point is 0.7070 p.u in DLG fault

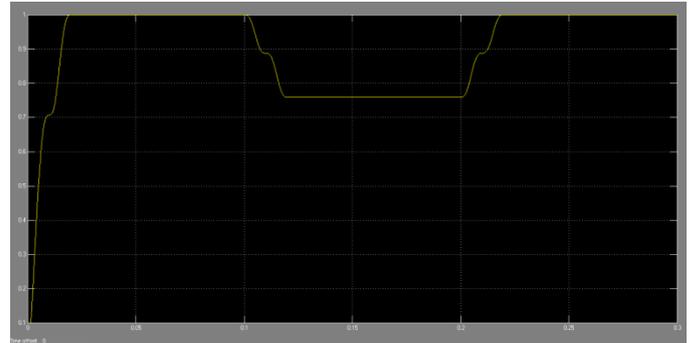


Fig. 9 Voltage at load point is 0.7587 p.u in LL fault

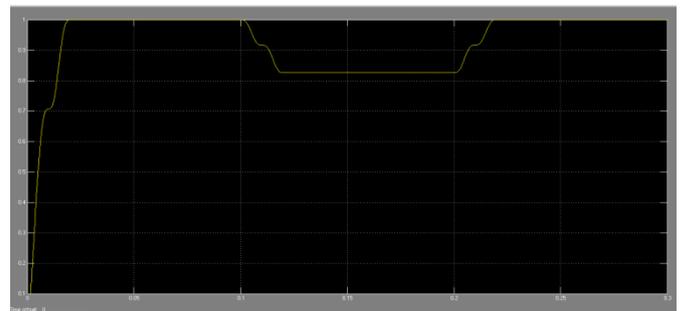


Fig. 10 voltage at load point is 0.8259 p.u in SLG fault

Figs. 7 to 10 show the simulation results of the test system for different types of fault without D-STATCOM. The fault occurs during when the fault resistance is 0.66 Ω

4.2 Simulation Model for the test system with insertion of D-STATCOM

To create distortion in the distribution system, different types of fault such as Three Phase to Ground (TPG), Double Line to Ground (DLG), Line to Line (LL), and Single Line to Ground (SLG) are injected.

Table 2 Results of voltage sags for different types of faults.

R_f in Ω	SLG fault	DLG fault	TLG fault	LL fault
0.86	0.9863	0.9858	0.9543	1.0152
0.76	0.9817	0.9806	0.9448	1.0143
0.66	0.9836	0.9801	0.9368	1.0169

Table 2 show the overall results of voltage sags in p.u with different types of fault. From the table, it can be observed that voltage sags improved with insertion of DSTATCOM. The value of voltage sags is between (0.9 to 1.02 p.u.)

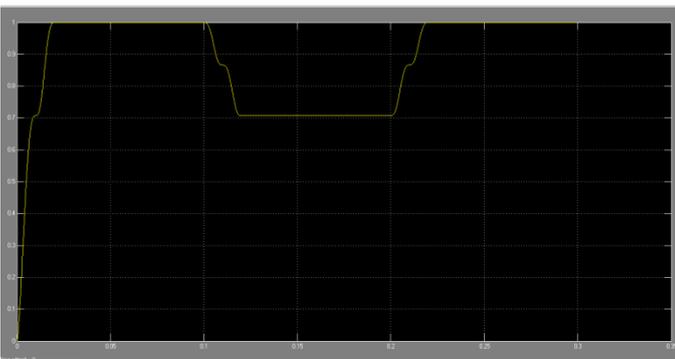


Fig. 7 Voltage at load point is 0.6600 p.u in TPG fault

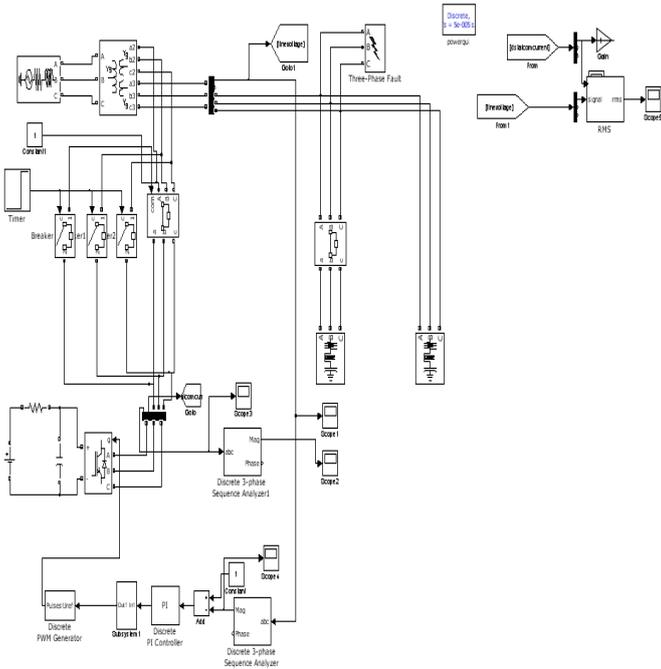


Fig. 11 Simulink Diagram for test system with D-STATCOM

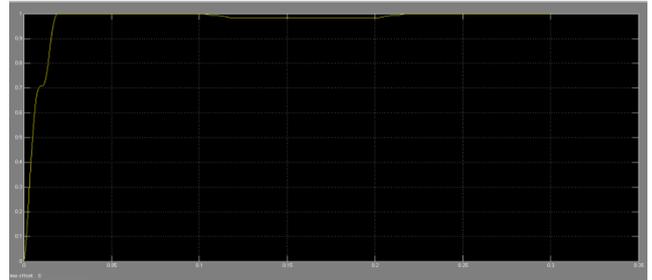


Fig. 15 voltage at load point is 0.9837 p.u in SLG fault

Figs. 12 to 15 show the simulation results of the test system for different types of fault with D-STATCOM. The fault occurs during when the fault resistance is 0.66 Ω.

Table 3 Results for different types fault before & after insert D-STATCOM when $R_f=0.86$

Types of Faults	Without D-STATCOM (p.u)	With D-STATCOM (p.u)	% of improvement
SLG	0.8679	0.9863	11.84
DLG	0.7833	0.9858	20.25
TPG	0.7515	0.9543	20.28
LL	0.8210	1.0152	19.42

From Table 3 it can be seen that with D-STATCOM the voltage sags has improved close to 1.0 p.u

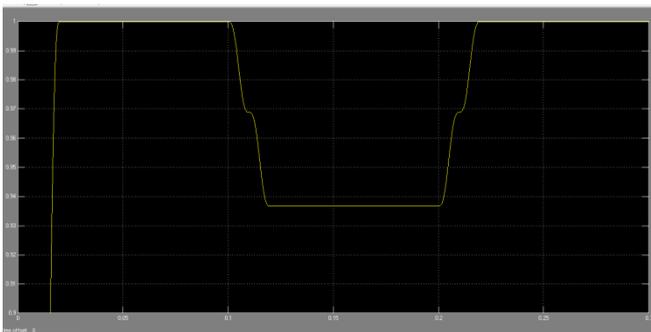


Fig. 12 Voltage at load point is 0.9367 p.u in TPG fault

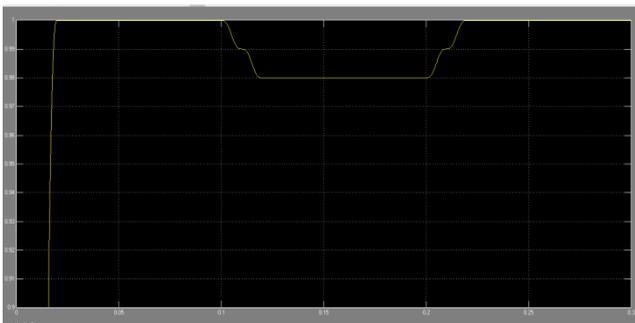


Fig. 13 Voltage at load point is 0.9888 p.u in DLG fault

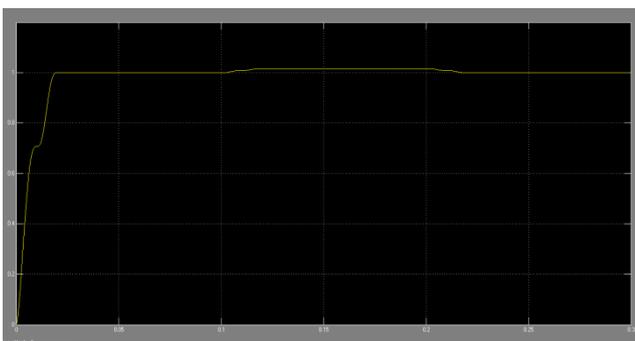


Fig. 14 voltage at load point is 1.0168 p.u in LL fault

4.3 Simulation Model for the test system with insertion of D-STATCOM with LCL Passive Filter

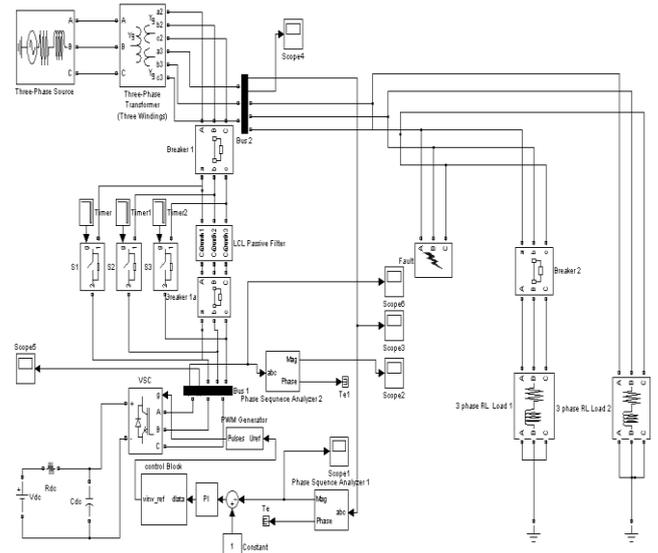


Fig. 16 Simulink Diagram for test system with insertion D-STATCOM with LCL Passive Filter

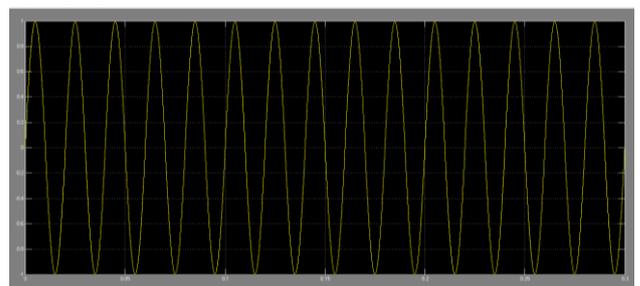


Fig. 17 waveform of output current with LCL Passive Filter

Fig. 17 shows the waveforms of output current. It is sinusoidal with LCL Passive filter was connected to the DSTATCOM. Figure 8.12 shows the spectrum of output current.

V. CONCLUSION

The simulation results show that the voltage sags can be mitigate by inserting D-STACTOM to the 11kV Radial distribution test system. By adding LCL passive filter to D-STACTOM, the Total Harmonic Distortions reduced and power factor can also increases close to unity. Thus, it can be concluded that by adding D-STACTOM with LCL passive filter the power quality is improved.

REFERENCES

- [1] N. G. Hingorani, "Power electronics in electrical utilities: role of power electronics in future power systems," *Proceedings of the IEEE* Vol. 76 No.4, pp.481-482, Apr. 1988
- [2] N. G. Hingorani and L. Gyugyi, "Understanding FACTS-concepts and technology of flexible AC transmission systems," *IEEE press, First Indian Edition*, 2001.
- [3] A.E. Hammad, Comparing the Voltage ource capability of Present and future Var Compensation Techniques in Transmission System, *IEEE Trans, on Power Delivery*. volume 1. No.1 Jan 1995.
- [4] G.Yalienkaya, M.H.J Bollen, P.A. Crossley, "Characterization of Voltage Sags in Industrial Distribution System", *IEEE transactions on industry applications*, volume 34, No. 4, July/August, PP.682-688, 1999
- [5] Haque, M.H., "Compensation Of Distribution Systems Voltage sags by DVR and D STATCOM", *Power Tech Proceedings, 2001IEEE Porto*, Volume 1, PP.10-13, September 2001.
- [6] Anaya-Lara O, Acha E., "Modeling and Analysis Of Custom Power Systems by SCAD/EMTDC", *IEEE Transactions on Power Delivery*, Volume 17, Issue: 2002, Pages: 266-272.
- [7] Bollen, M.H.J., "Voltage sags in Three Phase Systems", *Power Engineering Review*, IEEE, Volume 21, Issue :9, September 2001, PP: 11-15.

BIOGRAPHIES



Kollu Nageswara Rao was born in 1986. He graduated from Jawaharlal Nehru Technological University, Hyderabad in the year 2009. His Pursing P.G (PE&ED) in the Department of Electrical and Electronics Engineering at MIST, Sathupally, India. His research area includes Power Electronics and Drives.

His research areas include PWM techniques, DC to AC converters and control of electrical drives



Chalasani Hari Krishna was born in 1982. He graduated from Jawaharlal Nehru Technological University, Hyderabad in the year 2003. He received M.E degree from Satyabama University, Chennai in the year 2005. He presently Associate Professor in the Department of Electrical and Electronics Engineering at Mother Teresa Institute of Science and

Technology, India. His research area includes DTC and Drives. He presented 11 research papers in various national and international conferences and journals. His research areas include PWM techniques, DC to AC converters and control of electrical drives



Kiran Kumar Kuthadi recived his B.Tech degree in Electrical Engineering from Lakireddy Balireddy College of Engineering, Mylavaram, A.P and his M.Tech degree in Electrical Power Systems from J.N.T. University Anantapur, Anantapur, A.P in 2009. His special research of interest is its application to power system, FACTS

devices and their control. He is working as Asst. Professor in Sree Vahini Institute of Science & Technology, Tiruvuru. Andhra Pradesh. He presented 4 research papers in various national and international journals

Image Resolution Enhancement by Using Stationary and Discrete Wavelet Decomposition

Battula.R.V.S.Narayana¹ and K.Nirmala²

¹M.Tech Student, Department of ECE, Gudlavalleru Engineering College

²Assistant Professor, Department of ECE, Gudlavalleru Engineering College

Abstract: This work proposed an image resolution enhancement technique which is based on the interpolation of the high frequency subbands obtained by DWT. The proposed technique uses DWT to decompose an image into different subbands, and then the high frequency subband images have been interpolated. The interpolated high frequency subband coefficients have been corrected by using the high frequency subbands achieved by SWT of the input image. An original image is interpolated with half of the interpolation factor used for interpolation the high frequency subbands. Afterwards all these images have been combined using IDWT to generate a super resolved image. The proposed technique has been tested on well-known benchmark images, where their PSNR, Mean Square Error and Entropy results show the superiority of proposed technique over the conventional and state-of-art image resolution enhancement techniques.

Keywords: Discrete wavelet transforms, image super resolution, stationary wavelet transform.

I. INTRODUCTION

Discrete wavelet transform (DWT) [8] is one of the recent wavelet transforms used in image processing. DWT decomposes an image into different sub band images, namely low-low (LL), low- high (LH), high-low (HL), and high-high (HH). Another recent wavelet transform which has been used in several image processing applications is stationary wavelet transform (SWT) [9]. In short, SWT is similar to DWT but it does not use down-sampling, hence the sub bands will have the same size as the input image.

Resolution has been frequently referred as an important aspect of an image. Images are being processed in order to obtain more enhanced resolution. One of the commonly used techniques for image resolution enhancement is Interpolation.

The high frequency sub- bands obtained by SWT of the input image are being incremented into the interpolated high frequency subbands in order to correct the estimated coefficients. In parallel, the input image is also interpolated separately. Finally, corrected interpolated high frequency subbands and interpolated input image are combined by using inverse DWT (IDWT) to achieve a high resolution output image. In this work, we are proposing an image resolution enhancement technique which generates sharper high resolution image.

The proposed technique uses DWT to decompose a low resolution image into different subbands. Then the three high frequency subband images have been interpolated using bi-cubic interpolation. The proposed technique has been compared with conventional and state-of-art image resolution enhancement techniques. The conventional techniques used are the following:

- Interpolation techniques: bilinear interpolation and bicubic interpolation;
- wavelet zero padding (WZP). The state-of-art techniques used for comparison purposes are the following
- Regularity-preserving image interpolation [7];
- new edge-directed interpolation (NEDI) [10];
- hidden Markov model (HMM) [11];
- HMM-based image super resolution (HMM SR) [12];
- WZP and cycle-spinning (WZP-CS) [13];
- WZP, CS, and edge rectification (WZP-CS-ER) [14];
- DWT based super resolution (DWT SR) [15];
- Complex wavelet transform based super resolution (CWT SR) [5].

According to the quantitative and qualitative experimental results, the proposed technique over performs the aforementioned conventional and state-of-art techniques for image resolution enhancement.

II. PROPOSED IMAGE RESOLUTION ENHANCEMENT

In this correspondence, one level DWT (with Daubechies 9/7 as wavelet function) is used to decompose an input image into different subband images. Three high frequency subbands (LH, HL, and HH) contain the high frequency components of the input image. In the proposed technique, bi-cubic interpolation with enlargement factor of 2 is applied to high frequency subband images.

In image resolution enhancement by using interpolation the main loss is on its high frequency components (i.e., edges), which is due to the smoothing caused by interpolation. In order to increase the quality of the super resolved image, preserving the edges is essential.

In this work, DWT has been employed in order to preserve the high frequency components of the image. The redundancy and shift invariance of the DWT mean that DWT coefficients are inherently interpolable [9]. Downsampling in each of the DWT subbands causes information loss in the respective subbands.

That is why SWT is employed to minimize this loss. The interpolated high frequency subbands and the SWT high frequency subbands have the same size which means they can be added with each other.

The new corrected high frequency subbands can be interpolated further for higher enlargement. Also it is known that in the wavelet domain, the low resolution image is obtained by lowpass filtering of the high resolution image [16]. In other words, low frequency subband is the low resolution of the original image.

Therefore, instead of using low frequency subband, which contains less information than the original high resolution image, we are using the input image for the

interpolation of low frequency subband image. Using input image instead of low frequency subband increases the quality of the super resolved image. Fig. 1 illustrates the block diagram of the proposed image resolution enhancement technique.

By interpolating input image by $\alpha/2$, and high frequency subbands by 2 and α in the intermediate and final interpolation stages respectively, and then by applying IDWT, as illustrated in Fig. 1, the output image will contain sharper edges than the interpolated image obtained by interpolation of the input image directly.

This is due to the fact that, the interpolation of isolated high frequency components in high frequency subbands and using the corrections obtained by adding high frequency subbands of SWT of the input image, will preserve more high frequency components after the interpolation than interpolating input image directly.

III. RESULTS AND DISCUSSIONS

The below figure shows that super resolved image of Baboon's picture using proposed technique in (d) are much better than the low resolution image in (a), super resolved image by using the interpolation (b), and WZP (c). Note that the input low resolution images have been obtained by down-sampling the original high resolution images.

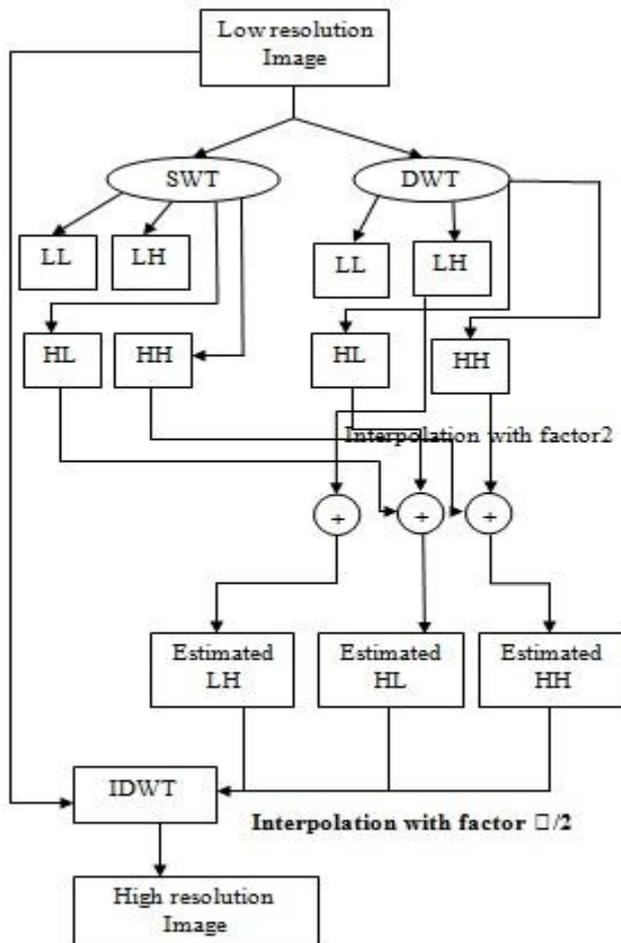


Figure-1. Block diagram of the proposed super resolution algorithm.

In order to show the effectiveness of the proposed method over the conventional and state-of-art image resolution enhancement techniques, four well-known test images

(Lena, Elaine, Baboon, and Peppers) with different features are used for comparison.

Table I compares the PSNR performance of the proposed technique using bicubic interpolation with conventional and state-of-art resolution enhancement techniques: bilinear, bicubic, WZP, NEDI, HMM, HMM SR, WZP-CS, WZP-CS-ER, DWT SR, CWT SR, and regularity-pre- serving image interpolation.

Additionally, in order to have more comprehensive comparison, the performance of the super resolved image by using SWT only (SWT-SR) is also included in the table. The results in Table I indicate that the proposed technique over-performs the aforementioned conventional and state-of-art image resolution enhancement techniques. Table I also indicates that the proposed technique over-performs the aforementioned conventional and state-of-art image resolution enhancement techniques.

Table-I. PSNR (DB) Resultsfor Resolution Enhancement from 128×128 to 512×512 of the Proposed Technique compared with the Conventional and State-Of-Art Image Resolution Enhancement Techniques

Techniques / Images	PSNR (dB)			
	Lena	Elaine	Baboon	Peppers
Bilinear	26.34	25.38	20.51	25.16
Bicubic	26.86	28.93	20.61	25.66
WZP(db, 9/7)	28.84	30.44	21.47	29.57
Regularity-Preserving Image Interpolation	28.81	29.97	21.18	28.52
NEDI	28.81	29.97	21.18	28.52
HMM	28.86	30.46	21.47	29.58
HMM SR	28.88	30.51	21.49	29.60
WZP-CS	29.27	30.78	21.54	29.87
WZP-CS-ER	29.36	30.89	21.56	30.05
DWT-SR	34.79	32.73	23.29	32.19
SWT-SR	32.01	31.25	22.74	29.46
Proposed Technique	34.82	35.01	23.87	33.06
Mean square error	4.925	4.5297	7.5819	5.7331
Entropy	7.392	7.2107	7.4451	6.9401

SWT high frequency subbands, and the input image. The proposed technique uses DWT to decompose an image into different subbands, and then the high frequency subband images have been interpolated. The interpolated high frequency subband coefficients have been corrected by using the high frequency subbands achieved by SWT of the input image. An original image is interpolated with half of the interpolation factor used for interpolation the high frequency subbands. After- wards all these images have been combined using IDWT to generate a super resolved imaged. The proposed technique has been tested on well-known benchmark images, where their PSNR and visual results show the superiority of proposed technique over the conventional and state-of-art image resolution enhancement techniques.

IV. CONCLUSION

This work proposed an image resolution enhancement technique based on the interpolation of the high frequency subbands obtained by DWT, correcting the high frequency subband estimation by using SWT high frequency subbands, and the input image. The proposed technique uses DWT to decompose an image into different subbands, and then the high frequency subband images have been interpolated. The interpolated high frequency subband coefficients have been corrected by using the high frequency subbands achieved by SWT of the input image. An original image is interpolated with half of the interpolation factor used for interpolation the high frequency subbands. Afterwards all these images have been combined using IDWT to generate a super resolved image. The proposed technique has been tested on well-known benchmark images, where their PSNR and visual results show the superiority of proposed technique over the conventional and state-of-art image resolution enhancement techniques.

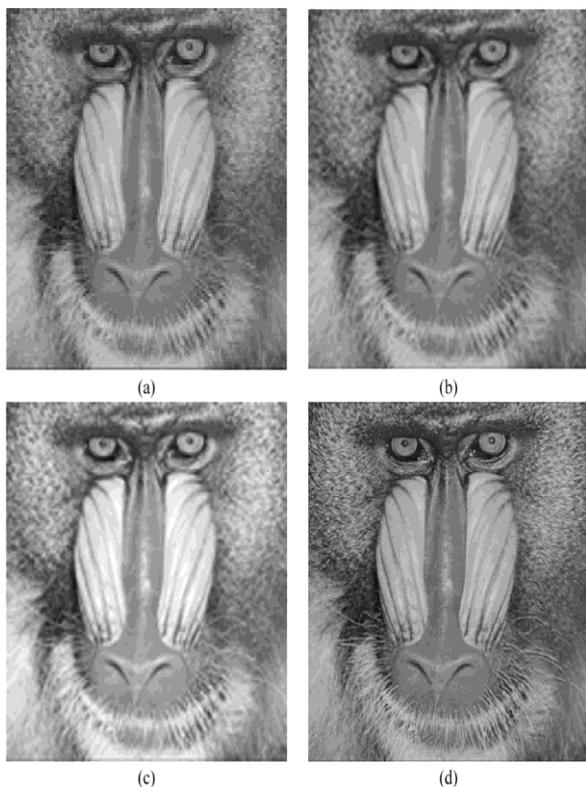


Figure-2. (a) Original low resolution Baboon's image. (b) Bicubic interpolated image. (c) Super resolved image using WZP. (d) Proposed technique.

REFERENCES

- 1) L. Yi-bo, X. Hong, and Z. Sen-yue, "The wrinkle generation method for facial reconstruction based on extraction of partition wrinkle line features and fractal interpolation," in *Proc. 4th Int. Conf. Image Graph.*, Aug. 22–24, 2007, pp. 933–937.
- 2) Y. Rener, J. Wei, and C. Ken, "Downsample-based multiple description coding and post-processing of decoding," in *Proc. 27th Chinese Control Conf.*, Jul. 16–18, 2008, pp. 253–256.
- 3) H. Demirel, G. Anbarjafari, and S. Izadpanahi, "Improved motion-based localized super resolution technique using discrete wavelet transform for low resolution video

enhancement," in *Proc. 17th Eur. Signal Process. Conf.*, Glasgow, Scotland, Aug. 2009, pp. 1097–1101.

- 4) Y. Piao, I. Shin, and H. W. Park, "Image resolution enhancement using inter-subband correlation in wavelet domain," in *Proc. Int. Conf. Image Process.*, 2007, vol. 1, pp. I-445–448.
- 5) H. Demirel and G. Anbarjafari, "Satellite image resolution enhancement using complex wavelet transform," *IEEE Geoscience and Remote Sensing Letter*, vol. 7, no. 1, pp. 123–126, Jan. 2010.
- 6) C. B. Atkins, C. A. Bouman, and J. P. Allebach, "Optimal image scaling using pixel classification," in *Proc. Int. Conf. Image Process.*, Oct. 7–10, 2001, vol. 3, pp. 864–867.
- 7) W.K.Carey,D.B.Chuang,andS.S.Hemami,"Regularity-preserving image interpolation," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp. 1295–1297, Sep. 1999.
- 8) S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York: Academic, 1999.
- 9) J. E. Fowler, "The redundant discrete wavelet transform and additive noise," Mississippi State ERC, Mississippi State University, Tech. Rep. MSSU-COE-ERC-04-04, Mar. 2004.
- 10) X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.
- 11) K. Kinebuchi, D. D. Muresan, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2001, vol. 3, pp. 7–11.
- 12) S. Zhao, H. Han, and S. Peng, "Wavelet domain HMT-based image super resolution," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2003, vol. 2, pp. 933–936.
- 13) A. Temizel and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning," *Electron. Lett.*, vol. 41, no. 3, pp. 119–121, Feb. 3, 2005.
- 14) A. Temizel and T. Vlachos, "Image resolution upscaling in the wavelet domain using directional cycle spinning," *J. Electron. Imag.*, vol. 14, no. 4, 2005.
- 15) G. Anbarjafari and H. Demirel, "Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image," *ETRI J.*, vol. 32, no. 3, pp. 390–394, Jun. 2010.
- 16) A. Temizel, "Image resolution enhancement using wavelet domain hidden Markov tree and coefficient sign estimation," in *Proc. Int. Conf. Image Process.*, 2007, vol. 5, pp. V-381–384.

AUTHORS

Mr. B.R.V.S. Narayana, completed his B.Tech in Electronics and Communication Engineering in 2010 and presently pursuing his M.Tech under the specialization of Digital Electronics and Communication Systems from Gudlavalleru Engineering College. His research interests include Image processing techniques, Logic design theory and CAD tools etc.

Ms. K. Nirmala, completed her M.Tech with distinction and presently working as Assistant Professor in Gudlavalleru Engineering College. She has an overall teaching experience of more than 2 years. Her research interests include Low Power VLSI, Image processing, etc.

Face Detection System On AdaBoost Algorithm Using Haar Classifiers

M.Gopi Krishna, A. Srinivasulu,

¹M. Tech Student, Department of Electronics (E.C.E), Vaagdevi Institute of Tech & Science, Peddasettipalli (Village), Proddatur (M.D), Kadapa (Dist), Andhra Pradesh A.P, India.

²M. Tech, Assistant Professor, Department of Electronics (E.C.E), Vaagdevi Institute of Tech & Science, Peddasettipalli (Village), Proddatur (M.D), Kadapa(Dist), Andhra Pradesh A.P, India.

Abstract: This paper presents an architecture for face detection based system on AdaBoost algorithm using Haar features. We describe here design techniques including image scaling, integral image generation, pipelined processing as well as classifier, and parallel processing multiple classifiers to accelerate the processing speed of the face detection system. Also we discuss the optimization of the proposed architecture which can be scalable for configurable devices with variable resources. The proposed architecture for face detection has been designed using Verilog HDL and implemented in Modelsim. Its performance has been measured and compared with an equivalent hardware implementation. We show about 35 time's increase of system performance over the equivalent hardware implementation.

Keywords: AdaBoost; architecture; face detection; Haar classifier; image processing; real-time.

I. INTRODUCTION

Face detection in image sequence has been an active research area in the computer vision field in recent years due to its potential applications such as monitoring and surveillance [1], human computer interfaces [2], smart rooms [3], intelligent robots [4], and biomedical image analysis [5]. Face detection is based on identifying and locating a human face in images regardless of size, position, and condition. Simple features such as color, motion, and texture are used for the face detection in early researches. However, these methods break down easily because of the complexity of the real world.

Face detection proposed by Viola and Jones [6] is most popular among the face detection approaches based on statistic methods. Although real-time face detection is possible using high performance computers, the resources of the system tend to be monopolized by face detection.

Therefore, this work has resulted in the development of a real-time face detection system employing an FPGA implemented system designed by Verilog HDL. Its performance has been measured and compared with an equivalent software implementation.

This paper is organized as follows: In Section II, we explain the face detection algorithm. In Section III, we describe the architecture, designed with Verilog HDL, of a face detection system using block diagrams. In Section IV, we show the implementation of the real-time face detection system in an FPGA and measure the corresponding performance. Finally, we conclude in Section V.

II. FACE DETECTION ALGORITHM

The face detection algorithm proposed by Viola and Jones is used as the basis of our design. The face detection

algorithm looks for specific Haar features of a human face. When one of these features is found, the algorithm allows the face candidate to pass to the next stage of detection. A face candidate is a rectangular section of the original image called a sub-window. Generally these sub-windows have a fixed size (typically 24×24 pixels).

This sub-window is often scaled in order to obtain a variety of different size faces. The algorithm scans the entire image with this window and denotes each respective section a face candidate [6].

A. Integral Image

The integral image is defined as the summation of the pixel values of the original image. The value at any location (x, y) of the integral image is the sum of the image's pixels above and to the left of location (x, y). "Fig. 1" illustrates the integral image generation.

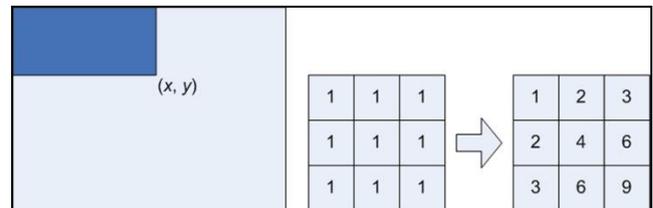


Figure 1. Integral image generation. The shaded region represents the sum of the pixels up to position (x, y) of the image. It shows a 3×3 image and its integral image representation.

B. Haar Features

Haar features are composed of either two or three rectangles. Face candidates are scanned and searched for Haar features of the current stage. The weights are constants generated by the learning algorithm. There are a variety of forms of features as seen below in "Fig. 2".

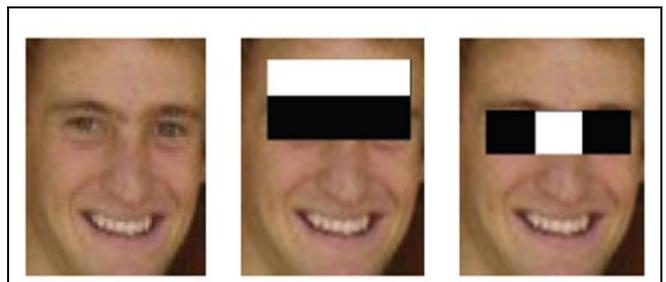


Figure 2. Examples of Haar features. Areas of white and black regions are multiplied by their respective weights and then summed in order to get the Haar feature value.

Each Haar feature has a value that is calculated by taking the area of each rectangle, multiplying each by their

respective weights, and then summing the results. The area of each rectangle is easily found using the integral image. The coordinate of the any corner of a rectangle can be used to get the sum of all the pixels above and to the left of that location using the integral image. Since L_1 is subtracted off twice it must be added back on to get the correct area of the rectangle. The area of the rectangle R , denoted as the rectangle integral, can be computed as follows using the locations of the integral image: $L_4 - L_3 - L_2 + L_1$.

B. Haar Feature Classifier

A Haar feature classifier uses the rectangle integral to calculate the value of a feature. The Haar feature classifier multiplies the weight of each rectangle by its area and the results are added together. Several Haar feature classifiers compose a stage. A stage comparator sums all the Haar feature classifier results in a stage and compares this summation with a stage threshold. The threshold is also a constant obtained from the AdaBoost algorithm. Each stage does not have a set number of Haar features. For example, Viola and Jones' data set used 2 features in the first stage and 10 in the second. All together they used a total of 38 stages and 6060 features [6]. Our data set is based on the OpenCV data set which used 22 stages and 2135 features in total.

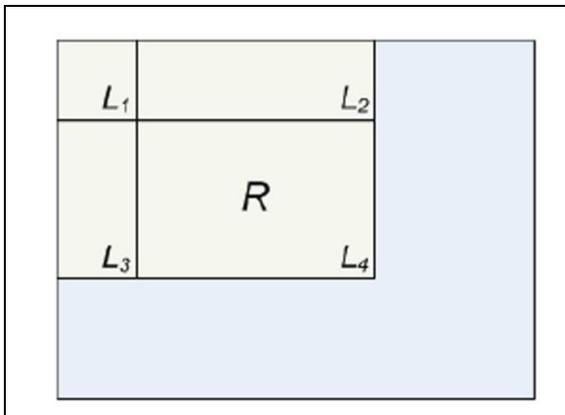


Figure 3. Calculating the area of a rectangle R is done using the corner of the rectangle: $L_4 - L_3 - L_2 + L_1$.

C. Cascade

The Viola and Jones face detection algorithm eliminates face candidates quickly using a cascade of stages. The cascade eliminates candidates by making stricter requirements in each stage with later stages being much more difficult for a candidate to pass. Candidates exit the cascade if they pass all stages or fail any stage. A face is detected if a candidate passes all stages. This process is shown in "Fig. 4".

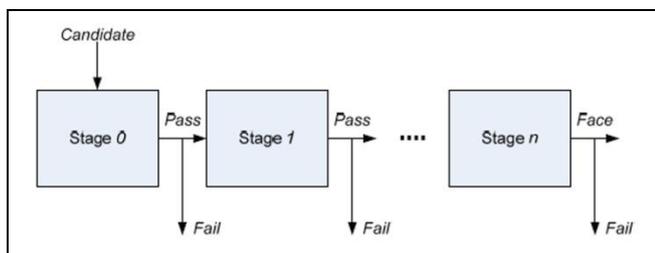


Figure 4. Cascade of stages. Candidate must pass all stages in the cascade to be concluded as a face.

III. IMPLEMENTATION

A. System Overview

We proposed architecture for a real-time face detection system. "Fig. 5" shows the overview of the proposed architecture for face detection. It consists of five modules: variant pose, illumination condition, Facial Expression, Occlusion, Uncontrolled Background, display. Face Detection systems are not only detected faces on uniform environment. In reality, Peoples are always located on complex background with different texture and object. These 'thing' are the major factors to affect the performance of face detection system

Face Detection System Architecture

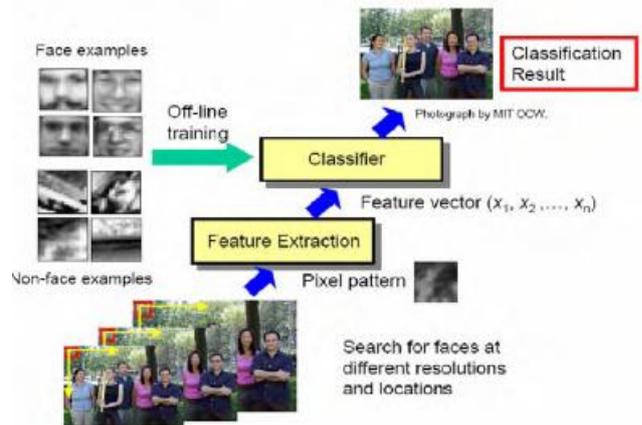


Figure 5. Block diagram of proposed face detection system.

B. Architecture for Face Detection

1) Variant Pose

Variant pose is occurred because of peoples not always orient to camera.

The image sync signal and the color image data are transferred from the image interface module. The image cropper crops the images based on the sync signals.

2) Illumination Condition

Different lighting and the quality of camera directly affect the quality of face. Sometimes it can be varied greater than facial expression and occlusion.

3) Facial Expression

The integral image generation requires substantial computation. A general purpose computer of Von Neumann architecture has to access image memory at least width×height times to get the value of each pixel when it processes an image with width×height pixels. For the incoming pixel where the coordinate is (x, y), the image line buffer controller.

4) Occlusion

Face detection not only deals with different faces, however, it need deal with any optional object. E.g. Hairstyle, sunglasses are the example of occlusion in face detection.

C. Integral Image

For the incoming pixel where the coordinate is (x, y), the image line buffer controller performs operations such as in "(1)", where n is the image window row size, p(x, y) is the incoming pixel value, and L(x, y) represents each pixel in the image line buffer.

$$L(x, y - k) = L(x, y - (k - 1)), \text{ where } 1 \leq k \leq n - 2 \quad (1)$$

$$L(x, y - k) = p(x, y), \text{ where } k = 0$$

The image window buffer stores pixel values moving from the image line buffer and its controller generates control signals for moving and storing the pixel values.

$$I(i - k, j) = I(i - (k - 1), j), \text{ where } 1 \leq k \leq m - 1 \quad (2)$$

$$I(i, j - l) = L(x, y - (l - 1)), \text{ where } 1 \leq l \leq n - 1$$

$$I(i - k, j - l) = p(i, j) = p(x, y), \text{ where } k = l = 0,$$

when $k + l = n - 1, 1 \leq k \leq n - 1, 0 \leq l \leq n - 2, m = 2n,$

$$I(i - k, j - l) = I(i - (k - 1), j - l) + I(i - (k - 1), j - (l + 1))$$

For the incoming pixel with coordinate (x, y), the image window buffer controller performs operation as in “(2)” where n and m are the row and column size of the image window buffer, respectively. p(i, j) is the incoming pixel value in the image window buffer; p(x, y) is the incoming pixel value; I(i, j) represents each of the pixels in the image window buffer; and L(x, y) represents each of the pixels in the image line buffer.

$$II(s - u, t - v) = II(s - u, t - v) + I(i - k, j - l) - I(i - (2n - 1), j - l) \quad (3)$$

where $0 \leq u \leq n - 1, 0 \leq v \leq n - 1, n - 1 \leq k \leq 2n - 2, 0 \leq l \leq n - 1$

Since pixels of an integral image window buffer are stored in registers, it is possible to access all integral pixels in the integral image window buffer simultaneously to perform the Haar feature classification.

“Fig. 6” shows all of the actions in the proposed architecture to generate the integral image. For every image from the frame grabber module, the integral image window buffer is calculated to perform the feature classification using the integral image.

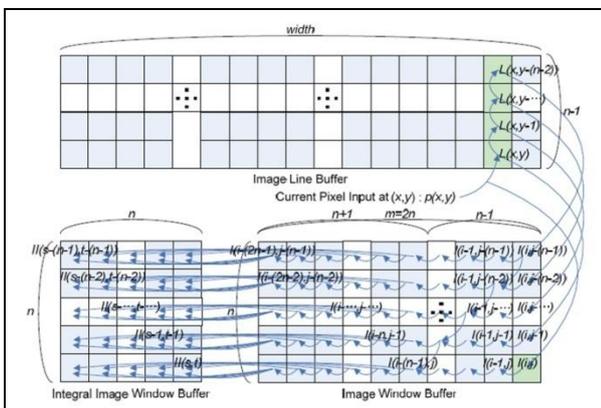


Figure 6. Architecture for generating integral image window.

A Haar classifier consists of two or three rectangles and their weight values, feature threshold value, and left and right values. Each rectangle presents four points using the coordinates (x, y) of most left and up point, width w, and height h as shown in “Fig. 7”.

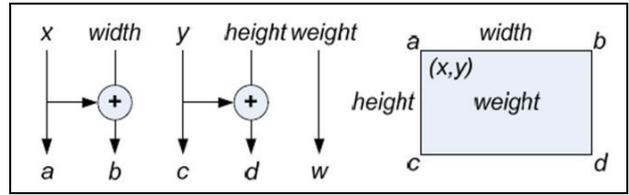


Figure 7. Rectangle calculation of Haar feature classifier. The integral pixel value of each rectangle can be calculated using these points from the integral image window buffer as shown in “Fig. 8”.

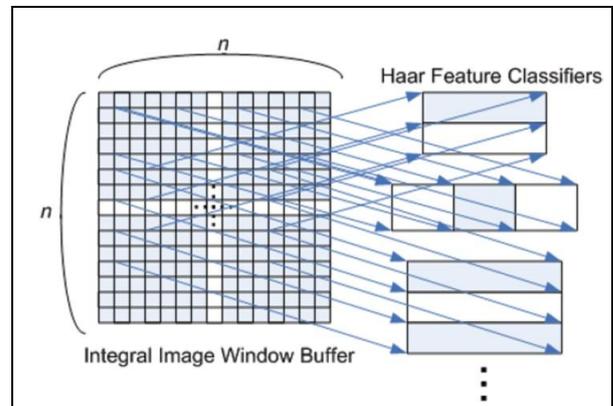


Figure 8. Simultaneous access to integral image window in order to calculate integral image of Haar feature classifiers. Four points of the rectangles of the Haar feature classifier are calculated by the method as shown in “Fig. 7”. The integral image values of Haar classifier are obtained from the integral image window buffer as shown in “Fig. 8”. Integral image value of each rectangle multiplies with its weight.

Display:

In the display module, the Digital Visual Interface (DVI) specification is applied to display the image sequence to the LCD monitor through a DVI transmitter in the DVI interface module. This module generates the sync signals and image data for the DVI transmitter.

Implementation

The proposed architecture for face detection has been designed using Verilog HDL and implemented in MODEL-SIM Altera 6.3 . We use the Haar feature training data from OpenCV to detect the frontal human faces based on the Viola and Jones algorithm. are 20x20, that includes a total of 22 stages, 2135 Haar classifiers, and 4630 Haar features.

1. Preprocessing

“System input is color images which included images of human faces or not, output is the human faces which is extracted from original images. In order to get the better result of detection, pre-processing is essential.

Gray scale Conversion

For getting to reduce the information of images, image should be done a converting to grayscale. Each color images (RGB images) are composed of 3 channels to present red, green and blue components in RGB space.

1. Given example images (R1,G1,B1),...,(Rn,Gn,Bn) where R, G, B are the value of red, green and blue respectively and ‘n’ is total number of pixel in given image.

2. The new grayscale images has pixel from G1,...,Gn, where using formula is as follows: $0.21R + 0.71G + 0.07 B = G$. Unlike averages method, this form is considering the ratio because of human perception.

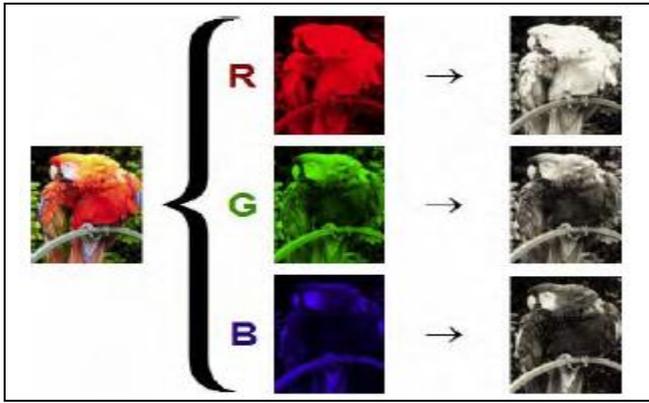
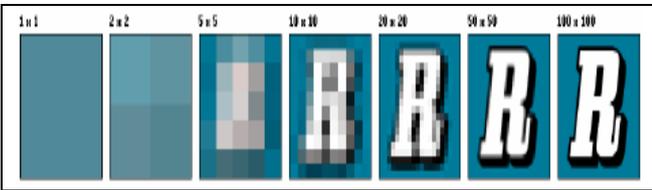


Figure 9 : RGB to Gray Scale Conversion

Image resizing

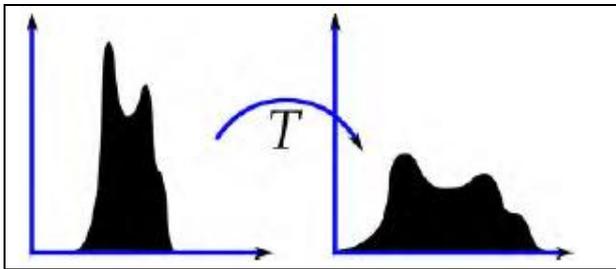
Images are synthesized by numerous of pixel which is the small unit in the image. For example, '0' is white and '255' is black in gray scale images.

Image has 3000 pixels in width and 2000 pixels in height which means it has 3000 x 2000 = 6,000,000 pixels or 6 megapixels. If the image has been resized into 1000 pixels in width and 600 pixels in height, image only has 0.6 megapixels. At least system only uses 1/10 timing to handle it.

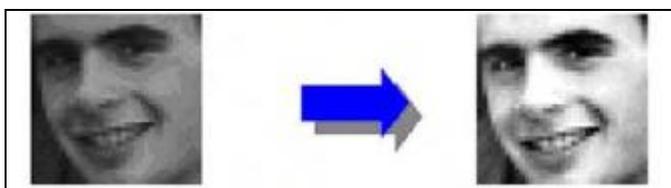


3. Histogram Equalization

Histogram equalization is a statistical method of images processing. It works as a statistical histogram of color distribution of the average scattered in the histogram. The change of the histogram after perform histogram equalization.



In the above chart, the shape of graph has been widened which is the meaning of average scattered in the histogram. This method usually increases the contrast of the input images. In face detection system, the left-hand side of below images is resized grayscale images.



Example of the process of histogram equalization

Algorithms of Histogram equalization:

1. Grayscale images has X_n pixels with i represent a value of gray level in each pixel. The following chart is represent the relationship between probability of occurrence and the value of each pixel:

Chart of Probability Density Function (PDF) And

$$\sum_{i=0}^{255} P(x_i) = 1$$

P_x is being histogram of images and normalized to [0,1]

2. Let us define the cumulative distribution function as follows:

$$F'(z) = \sum_{i=0}^z P(x_i) \quad z = 0,1,2...255$$

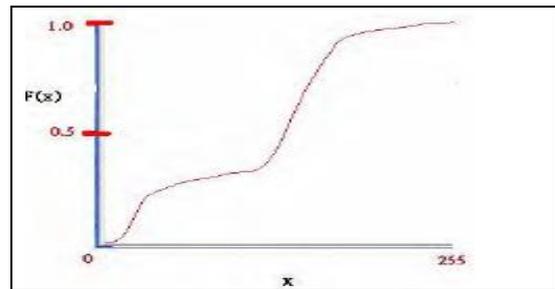


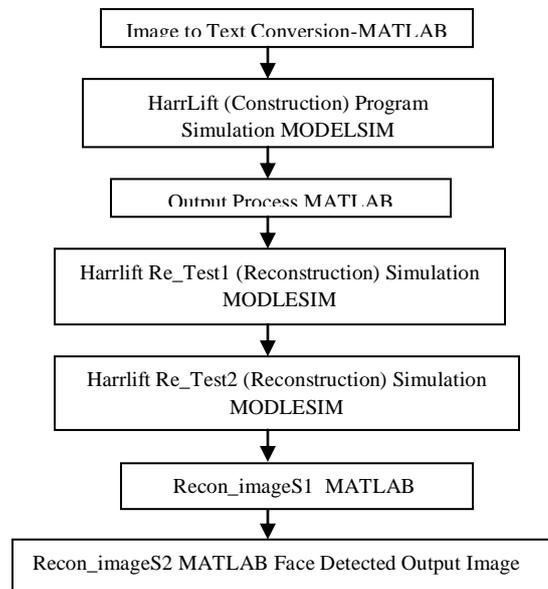
Chart of Cumulative distribution function

3. Minimum and maximum value are found and applied into following equation to find out the histogram equalization of each pixel:

$$h(v) = \text{round} \left(\frac{cdf(v) - cdf_{min}}{(M \times N) - cdf_{min}} \times (L - 1) \right)$$

Where cdf_{min} is the minimum value of CDF, M is the width of image and N is the height of image. L represent a large value of grey level, = 25

SIMULATION FLOW



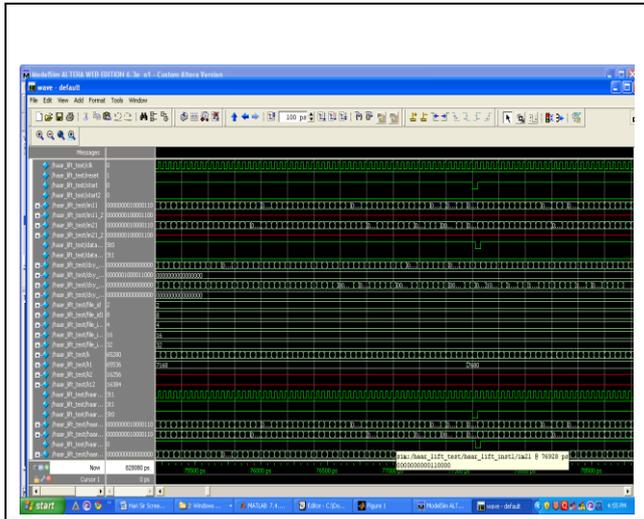


FIGURE: WAVE FORMS IN MODELSIM

IV. EXPERIMENTS/RESULTS

A high frame processing rate and low latency are important for many applications. We measure the performance of the proposed architecture for the face detection system. Face detection system when it is applied to a camera, which produces images consisting of 640×480 pixels at 60 frames per second.



Figure 10. Experimental result of face detection system.

V. CONCLUSION

We present face detection based on the AdaBoost algorithm using Haar features. In our architecture, the scaling image technique is used instead of the scaling sub-window, and the integral image window is generated instead of the integral image contains whole image during one clock cycle.

The Haar classifier is designed using a pipelined scheme, and the triple classifier which three single classifiers processed in parallel is adopted to accelerate the processing speed of the face detection system.

Finally, the proposed architecture is implemented on a Modelsim Altera 6.3 and its performance is measured and compared with an equivalent hardware implementation. We show about 35 times increase of system performance over

the equivalent software implementation. We plan to implement more classifiers to improve our design.

We have demonstrated that this face detection, combined with other technologies, can produce effective and powerful applications.

REFERENCES

- [1] Z. Guo, H. Liu, Q. Wang and J. Yang, "A Fast Algorithm of Face Detection for Driver Monitoring," In Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications, vol. 2, pp.267 - 271, 2006.
- [2] M. Yang, N. Ahuja, "Face Detection and Gesture Recognition for Human-Computer Interaction," The International Series in Video Computing, vol.1, Springer, 2001.
- [3] Z. Zhang, G. Potamianos, M. Liu, T. Huang, "Robust Multi- View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming," International Conference on Automatic Face and Gesture Recognition, pp.407-412, 2006.
- [4] W. Yun; D. Kim; H. Yoon, "Fast Group Verification System for Intelligent Robot Service," IEEE Transactions on Consumer Electronics, vol.53, no.4, pp.1731-1735, Nov. 2007.
- [5] V. Ayala-Ramirez, R. E. Sanchez-Yanez and F. J. Montecillo-Puente "On the Application of Robotic Vision Methods to Biomedical Image Analysis," IFMBE Proceedings of Latin American Congress on Biomedical Engineering, pp.1160-1162, 2007.
- [6] P. Viola and M. Jones, "Robust real-time object detection," International Journal of Computer Vision, 57(2), 137-154, 2004.
- [7] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," Journal of Computer and System Sciences, no. 55, pp. 119-139, 1997.

Mean Estimation with Imputation in Two-Phase Sampling

Narendra Singh Thakur, Kalpana Yadav, Sharad Pathak*

Center for Mathematical Sciences (CMS), Banasthali University, Banasthali, Rajasthan
*Department of Mathematics and Statistics, Dr. H. S. Gour Central University, Sagar (M.P.)

Abstract: Missing data is a problem encountered in almost every data collection activity but particularly in sample survey. The missing data naturally occurs in sample surveys when some, not all sampling units refuse or unable to participate in the survey or when data for specific items on a questionnaire completed for an otherwise cooperating unit are missing. Imputation is a methodology, which uses available data as a tool for the replacement of missing observations. Imputation methods used to fill the non responses and lead, under definite conditions, to suitable inference. This article suggests some imputation methods and discusses the properties of their mean estimators. Numerical study is performed over two populations using the expressions of bias and m.s.e and efficiency compared with existing estimators.

I. Introduction

In literature, several imputation techniques are described, some of them are better over others. Rubin (1976) addressed three concepts: OAR (observed at random), MAR (missing at random), and PD (parametric distribution). He defined that if the probability of the observed missingness pattern, given the observed and unobserved data, does not depend on the value of the unobserved data, then data are MAR. The observed data are observed at random (OAR) if for each possible value of the missing data and the parameter ϕ the conditional probability of the observed pattern of missing data given the missing data and the observed data, is the same for all possible values of the observed data. Heitzen and Basu (1996) have distinguished the meaning of MAR and MCAR in a very nice way. In what follows MCAR (missing completely at random) is used.

Little and Rubin (1987) define three different classes of missingness. They defined the key terms used in discussing missingness in the literature. Data missing on Y are observed at random (OAR) if missingness on Y is not a function of X . Phrased another way, if X determines missingness on Y , the data are not OAR. Data missing on Y are missing at random (MAR) if missingness on Y is not a function of Y . Phrased another way, if Y determines missingness on Y , the data are not MAR. Data are Missing Completely at Random (MCAR) if missingness on Y is unrelated to X or Y . In other words $MCAR = OAR + MAR$. If the data are MCAR or at least MAR, then the missing data mechanism is considered "ignorable." Otherwise, the missing data mechanism is considered "non-ignorable."

There are different ways and means to control non-response. One way of dealing with the problem of non-response is to make more efforts to collect information by taking a sub-sample of units not responding at the first attempt. Another way of dealing with the problem of non-response is to estimate the probability of responding informants of their being at home at a specified point of time and weighting results with the inverse of this probability. A technique to deal with the problem of non-response was developed by Hansen and Hurwitz (1946). They assumed that the population is divided into two classes, a *response class* who respond in the first attempt and a *non-response class* who did not.

A questionnaire contains many questions that we call items. When item non-response occurs, substantial information about the non-respondent is usually available from other items on the questionnaire. Many imputation methods in literature use selection of these items as auxiliary variable in assigning values to the i^{th} non-respondent for item y . Rao and Sitter (1995), Singh and Horn (2000), Ahmed et al. (2006) and Shukla and Thakur (2008) have given applications of various imputation procedures.

Let the variable Y is of main interest and X be an auxiliary variable correlated with Y and the population mean \bar{X} of auxiliary variable is unknown. A large preliminary simple random sample (without replacement) S' of n' units is drawn from the population $\Omega = (1, 2, \dots, N)$ to estimate \bar{X} and a secondary sample S of size n ($n < n'$) drawn as a sub-sample of the sample S' to estimate the population mean of main variable. Let the sample S contains n_1 responding units and $n_2 = (n - n_1)$ non-responding units. Using the concept of post-stratification, sample may be divided into two groups: responding (R_1) and non-responding (R_2).

The sample may be considered as stratified into two classes namely a *response class* and *non-response class*, then the procedure is known as *post-stratification*. Sukhatme (1984) advocates that post-stratification procedure is as precise as the stratified sampling under proportional allocation if the sample size is large enough. Estimation problem in sample surveys, in the setup of post-stratification, under non-response situation is studied due to Shukla and Dubey (2004 and 2008). Shukla et al. (2009) have also given the concept of utilization of \bar{X}_2 (population mean of non-response group of X) in imputation for missing observations of auxiliary information due to non-response.

Now it may be consider the population has two types of individuals like N_1 as number of respondents (R_1) and N_2 non-respondents (R_2), Thus the total N units of the population will comprise N_1 and N_2 , respectively, such that $N = N_1 + N_2$. The population proportions of units in the R_1 and R_2 groups are expressed as $W_1 = N_1/N$ and $W_2 = N_2/N$ such that $W_1 + W_2 = 1$. Further, let \bar{Y} and \bar{X} be the population means of Y and X respectively. For every unit $i \in R_1$, the value y_i is observed

available. However, for the units $i \in R_2$, the y_i 's are missing and imputed values are to be derived. The i^{th} value x_i of auxiliary variate is used as a source of imputation for missing data when $i \in R_2$. This is to assume that for sample S , the data $x_s = \{x_i : i \in S\}$ are known. The following notations are used in this paper:

\bar{x}_n, \bar{y}_n : the sample mean of X and Y respectively in S ; \bar{x}_1, \bar{y}_1 : the sample mean of X and Y respectively in R_1 ;

S_x^2, S_y^2 : the population mean squares of X and Y respectively; C_x, C_y : the coefficient of variation of X and Y respectively; ρ : Correlation Coefficient in population between X and Y respectively.

Further, consider few more symbolic representations:

$$L = E\left(\frac{1}{n_i}\right) = \left[\frac{1}{nW_1} + \frac{(N-n)(1-W_1)}{(N-1)n^2W_1^2} \right], \quad M = \frac{(N-n)(n-n_1)n'N}{nn_1^2(N-1)(N-n')}, \quad Q = \frac{nn_1^2(N-n')(N-1)}{n'N(N-n)(n-n_1) - 2nn_1^2(N-n')(N-1)}.$$

Under this setup as describe above in case of simple random sampling without replacement and assuming \bar{X} is known, some well known imputation methods are given below:

1.1. Mean Method of Imputation:

For y_i define $y_{.i}$ as
$$y_{.i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \bar{y}_r & \text{if } i \in R_2 \end{cases} \quad \dots(1.1)$$

Using above, the imputation-based estimator of population mean \bar{Y} is:
$$\bar{y}_m = \frac{1}{r} \sum_{i \in R} y_{.i} = \bar{y}_r \quad \dots(1.2)$$

Lemma 1.1: The bias and mean squared error is given by: (i) $B(\bar{y}_m) = 0 \quad \dots(1.3)$

(ii)
$$V(\bar{y}_m) \approx \left(\frac{1}{r} - \frac{1}{N}\right) S_y^2 \quad \dots(1.4)$$

1.2. Ratio Method of Imputation:

For y_i and x_i , define $y_{.i}$ as:
$$y_{.i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \hat{b}x_i & \text{if } i \in R_2 \end{cases} \quad \text{where } \hat{b} = \frac{\sum_{i \in R} y_i}{\sum_{i \in R} x_i} \quad \dots(1.5)$$

Under this, the imputation-based estimator is:
$$\bar{y}_s = \frac{1}{n} \sum_{i \in S} y_{.i} = \bar{y}_r \left(\frac{\bar{x}_n}{\bar{x}_r} \right) = \bar{y}_{RAT} \quad \dots(1.6)$$

where $\bar{y}_r = \frac{1}{r} \sum_{i \in R} y_i$, $\bar{x}_r = \frac{1}{r} \sum_{i \in R} x_i$ and $\bar{x}_n = \frac{1}{n} \sum_{i \in S} x_i$

Lemma 1.2: The bias and mean squared error of \bar{y}_{RAT} is given by: (i) $B(\bar{y}_{RAT}) = \bar{Y} \left(\frac{1}{r} - \frac{1}{n} \right) (C_x^2 - \rho C_y C_x) \quad \dots(1.7)$

(ii)
$$M(\bar{y}_{RAT}) \approx \left(\frac{1}{n} - \frac{1}{N}\right) S_y^2 + \left(\frac{1}{r} - \frac{1}{n}\right) [S_y^2 + R_1^2 S_x^2 - 2R_1 S_{xy}] \quad \text{where } R_1 = \frac{\bar{Y}}{\bar{X}} \quad \dots(1.8)$$

1.3. Compromised Method of Imputation:

Singh and Horn (2000) proposed Compromised imputation procedure as given below:

$$y_{.i} = \begin{cases} (an/r)y_i + (1-\alpha)\hat{b}x_i & \text{if } i \in R_1 \\ (1-\alpha)\hat{b}x_i & \text{if } i \in R_2 \end{cases} \quad \dots(1.9)$$

where α is a suitably chosen constant, such that the resultant variance of the estimator is optimum. The imputation-based estimator, for this case, Estimator of population mean is
$$\bar{y}_{COMP} = \left[\alpha \bar{y}_r + (1-\alpha) \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r} \right] \quad \dots(1.10)$$

Lemma 1.3: The bias, mean squared error and minimum mean squared error at $\alpha = 1 - \rho \frac{C_y}{C_x}$ of \bar{y}_{COMP} is given by

(i)
$$B(\bar{y}_{COMP}) = \bar{Y} (1-\alpha) \left(\frac{1}{r} - \frac{1}{n} \right) (C_x^2 - \rho C_y C_x) \quad \dots(1.11)$$

$$(ii) M(\bar{y}_{comp}) \approx \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) S_y^2 + \left(\frac{1}{r} - \frac{1}{n} \right) [S_y^2 + R_1^2 S_z^2 - 2R_1 S_{yz}] \right\} - \left(\frac{1}{r} - \frac{1}{n} \right) \alpha^2 \bar{Y}^2 C_x^2 \quad \dots(1.12)$$

$$(iii) M(\bar{y}_{com})_{min} = \left[\left(\frac{1}{r} - \frac{1}{N} \right) - \left(\frac{1}{r} - \frac{1}{n} \right) \rho^2 \right] S_y^2 \quad \dots(1.13)$$

1.4. Ahmed's Methods:

$$(A) y_{7i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \bar{y}_r + \frac{nk_1}{(n-r)} (\bar{X} - \bar{x}) + k_2 (x_i - \bar{x}_r) & \text{if } i \in R_2 \end{cases} \quad \dots(1.14)$$

Under this method, the point estimator of \bar{Y} is: $t_7 = \bar{y}_r + k_1 (\bar{X} - \bar{x}) + k_2 (\bar{x} - \bar{x}_r)$... (1.15)

Lemma 1.4: The bias, variance and minimum variance at $k_1 = k_2 = \frac{S_{xy}}{S_x^2}$ of t_7 is given by: (i) $B[t_7] = 0$... (1.16)

$$(ii) V(t_7) = \left(\frac{1}{r} - \frac{1}{N} \right) S_y^2 - 2S_{xy} \left[k_1 \left(\frac{1}{n} - \frac{1}{N} \right) + k_2 \left(\frac{1}{r} - \frac{1}{n} \right) \right] + S_x^2 \left[k_1^2 \left(\frac{1}{n} - \frac{1}{N} \right) + k_2^2 \left(\frac{1}{r} - \frac{1}{n} \right) \right] \quad \dots(1.17)$$

$$(iii) V(t_7)_{min} = \left(\frac{1}{r} - \frac{1}{N} \right) S_y^2 (1 - \rho^2) \quad \dots(1.18)$$

$$(B) y_{8i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \left[\frac{\bar{y}_r \left(x_i + \frac{r}{n-r} \bar{x}_r \right)}{\theta_1 \bar{x}_r + (1-\theta_1) \bar{x}} - \frac{r}{n-r} \bar{y}_r \right] & \text{if } i \in R_2 \end{cases} \quad \dots(1.19)$$

under this setup, the point estimator of \bar{Y} is: $t_8 = \frac{\bar{y}_r \bar{x}}{\theta_1 \bar{x}_r + (1-\theta_1) \bar{x}}$... (1.20)

Lemma 1.5: The bias, mean squared error and minimum mean squared error at $\theta_1 = \rho \frac{C_y}{C_x}$ of t_8 is given by

$$(i) B(t_8) \approx \left(\frac{1}{r} - \frac{1}{n} \right) \bar{Y} (\theta_1^2 C_x^2 - \theta_1 \rho C_y C_x) \quad \dots(1.21)$$

$$(ii) M(t_8) \approx \bar{Y}^2 \left[\left(\frac{1}{r} - \frac{1}{N} \right) C_y^2 + \theta_1^2 \left(\frac{1}{r} - \frac{1}{n} \right) C_x^2 - 2\theta_1 \left(\frac{1}{r} - \frac{1}{n} \right) \rho C_y C_x \right] \quad \dots(1.22)$$

$$(iii) M(t_8)_{min} \approx \left(\frac{1}{r} - \frac{1}{N} \right) S_y^2 - \left(\frac{1}{r} - \frac{1}{n} \right) \frac{S_{xy}^2}{S_x^2} \quad \dots(1.23)$$

$$(C) y_{9i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \frac{1}{(n-r)} \left[\frac{n \bar{y}_r \bar{X}}{\theta_2 \bar{x} + (1-\theta_2) \bar{X}} - r \bar{y}_r \right] & \text{if } i \in R_2 \end{cases} \quad \dots(1.24)$$

Under this method, the point estimator of \bar{Y} is: $t_9 = \frac{\bar{y}_r \bar{X}}{\theta_2 \bar{x} + (1-\theta_2) \bar{X}}$... (1.25)

Lemma 1.6: The bias, mean squared error and minimum mean squared error at $\theta_2 = \rho \frac{C_y}{C_x}$ of t_8 is given by

$$(i) B(t_9) \approx \left(\frac{1}{n} - \frac{1}{N} \right) \bar{Y} (\theta_2^2 C_x^2 - \theta_2 \rho C_y C_x) \quad \dots(1.26)$$

$$(ii) M(t_9) \approx \bar{Y}^2 \left[\left(\frac{1}{r} - \frac{1}{N} \right) C_y^2 + \theta_2^2 \left(\frac{1}{n} - \frac{1}{N} \right) C_x^2 - 2\theta_2 \left(\frac{1}{n} - \frac{1}{N} \right) \rho C_y C_x \right] \quad \dots(1.27)$$

$$(iii) M(t_9)_{\min} \approx \left(\frac{1}{r} - \frac{1}{N} \right) S_y^2 - \left(\frac{1}{n} - \frac{1}{N} \right) \frac{S_{xy}^2}{S_x^2} \quad \dots(1.28)$$

$$(d) y_{10i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \frac{1}{(n-r)} \left[\frac{n \bar{y}_r \bar{X}}{\theta_3 \bar{x} + (1-\theta_3) \bar{X}} - r \bar{y}_r \right] & \text{if } i \in R_2 \end{cases} \quad \dots(1.29)$$

under this, the point estimator of population mean \bar{Y} is: $t_{10} = \frac{\bar{y}_r \bar{X}}{\theta_3 \bar{x}_r + (1-\theta_3) \bar{X}}$... (1.30)

Lemma 1.7: The bias, mean squared error and minimum mean squared error at $\theta_3 = \rho \frac{C_y}{C_x}$ of t_{10} is given by

$$(i) B(t_{10}) \approx \left(\frac{1}{r} - \frac{1}{N} \right) \bar{Y} (\theta_3^2 C_x^2 - \theta_3 \rho C_y C_x) \quad \dots(1.31)$$

$$(ii) M(t_{10}) \approx \bar{Y}^2 \left(\frac{1}{r} - \frac{1}{N} \right) [C_y^2 + \theta_3^2 C_x^2 - 2\theta_3 \rho C_y C_x] \quad \dots(1.32)$$

$$(iii) M(t_{10})_{\min} = \left(\frac{1}{r} - \frac{1}{N} \right) S_y^2 (1 - \rho^2) \quad \dots(1.33)$$

II. Large Sample Approximations:

Let $\bar{y}_1 = \bar{Y}(1+e_1)$; $\bar{x}_1 = \bar{X}(1+e_2)$; $\bar{x}_n = \bar{X}(1+e_3)$ and $\bar{x}' = \bar{X}(1+e'_3)$, which implies the results $e_1 = \frac{\bar{y}_1}{\bar{Y}} - 1$; $e_2 = \frac{\bar{x}_1}{\bar{X}} - 1$; $e_3 = \frac{\bar{x}_n}{\bar{X}} - 1$ and $e'_3 = \frac{\bar{x}'}{\bar{X}} - 1$. Now by using the concept of two-phase sampling and the mechanism of MCAR, for given n_1 , n and n' [see Rao and Sitter (1995)] we have:

$$E(e_1) = E[E(e_1)|n_1] = E\left[\left(\frac{\bar{y}_1 - \bar{Y}}{\bar{Y}}\right) \middle| n_1\right] = \frac{\bar{Y} - \bar{Y}}{\bar{Y}} = 0; \text{ Similarly, } E(e_2) = E(e_3) = E(e'_3) = 0;$$

$$E(e_1^2) = E\left[\left(\frac{\bar{y}_1 - \bar{Y}}{\bar{Y}}\right)^2 \middle| n_1\right] = \left(E\left(\frac{1}{n_1}\right) - \frac{1}{n}\right) C_y^2 = \left(L - \frac{1}{n'}\right) C_y^2; \quad E(e_2^2) = \left(L - \frac{1}{n'}\right) C_x^2; \quad E(e_3^2) = \left(\frac{1}{n} - \frac{1}{n'}\right) C_x^2;$$

$$E(e_1 e_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) C_x^2; \quad E(e_1 e_3) = E(e_1 e_2 / n_1) = E\left[\left(\frac{(\bar{y}_1 - \bar{Y})(\bar{x}_1 - \bar{X})}{\bar{Y} \bar{X}}\right) \middle| n_1\right] = \left(E\left(\frac{1}{n_1}\right) - \frac{1}{n}\right) \rho C_y C_x$$

$$= \left(L - \frac{1}{n'}\right) \rho C_y C_x; \quad E(e_2 e_3) = \left(\frac{1}{n} - \frac{1}{n'}\right) \rho C_y C_x; \quad E(e_1 e'_3) = \left(\frac{1}{n'} - \frac{1}{N}\right) \rho C_y C_x; \quad E(e_2 e_3) = \left(\frac{1}{n} - \frac{1}{n'}\right) C_x^2;$$

$$E(e_2 e'_3) = \left(\frac{1}{n'} - \frac{1}{N}\right) C_x^2; \quad E(e_3 e'_3) = \left(\frac{1}{n'} - \frac{1}{N}\right) C_x^2$$

III. Proposed Different Imputation Methods

Let y_{vji} denotes the i^{th} available observation for the j^{th} imputation. We suggest the following imputation methods:

$$(1) y_{v7i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \bar{y}_1 + \frac{1}{(1-W_1)} \left[h(\bar{x}' - \bar{x}_n) + (1-W_1)k(x_i - \bar{x}_1) \right] & \text{if } i \in R_2 \end{cases} \quad \dots(3.1)$$

where h and k are suitably chosen constants, such that the variance the resultant estimator is minimum. Under this

$$T_{v7} = \bar{y}_1 + h(\bar{x}' - \bar{x}_n) + k(\bar{x}_n - \bar{x}_1) \quad \dots(3.2)$$

$$(2) y_{v8i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \frac{\bar{y}_1}{(1-W_1)} \left[\frac{(x_i(1-W_1) + W_1 \bar{x}_1)}{\theta \bar{x}_1 + (1-\theta) \bar{x}_n} - W_1 \right] & \text{if } i \in R_2 \end{cases} \quad \dots(3.3)$$

where θ is suitably chosen constant, such that the variance the resultant estimator is minimum. Under this

$$T_{v8} = \frac{\bar{y}_1 \bar{x}_n}{\theta \bar{x}_1 + (1-\theta) \bar{x}_n} \quad \dots(3.4)$$

$$(3) y_{v9i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \frac{\bar{y}_1}{(1-W_1)} \left[\frac{\bar{x}_1}{\phi \bar{x}_n + (1-\phi) \bar{x}_1} - W_1 \right] & \text{if } i \in R_2 \end{cases} \quad \dots(3.5)$$

where ϕ is suitably chosen constant, such that the variance the resultant estimator is minimum.

Under this, the point estimator of population mean \bar{Y} is

$$T_{v9} = \frac{\bar{y}_1 \bar{x}}{\phi \bar{x}_n + (1-\phi) \bar{x}} \quad \dots(3.6)$$

$$(4) y_{v10i} = \begin{cases} y_i & \text{if } i \in R_1 \\ \frac{\bar{y}_1}{(1-W_1)} \left[\frac{\bar{x}}{\psi \bar{x}_r + (1-\psi) \bar{x}} - W_1 \right] & \text{if } i \in R_2 \end{cases} \quad \dots(3.7)$$

where ψ is suitably chosen constant, such that the variance the resultant estimator is minimum. Under this, the point estimator of population mean \bar{Y} is

$$T_{v10} = \frac{\bar{y}_1 \bar{x}}{\psi \bar{x}_r + (1-\psi) \bar{x}} \quad \dots(3.8)$$

IV. Bias and M.S.E of Proposed Methods:

Let $B(\cdot)$ and $M(\cdot)$ denote the bias and mean squared error (*M.S.E.*) of an estimator under a given sampling design. The properties of estimators are derived in the following theorems respectively.

Theorem 4.1:

(1) The estimator T_{v7} in terms of e_1, e_2, e_3 and e_3' is:

$$T_{v7} = \bar{Y}(1 + e_1) + h\bar{X}(e_3' - e_3) + k\bar{X}(e_3 - e_2) \quad \dots(4.1)$$

Proof: $T_{v7} = \bar{y}_1 + h(\bar{x}' - \bar{x}_n) + k(\bar{x}_n - \bar{x}_1) = \bar{Y}(1 + e_1) + h\bar{X}(e_3' - e_3) + k\bar{X}(e_3 - e_2)$

(2) The estimator T_{v7} is unbiased i.e. $B[T_{v7}] = 0 \quad \dots(4.2)$

Proof: $B(T_{v7}) = E[T_{v7} - \bar{Y}] = \bar{Y} - \bar{Y} = 0$

(3) The variance of T_{v7} is

$$V(T_{v7}) = \left(L - \frac{1}{n'} \right) S_y^2 + \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) (h^2 S_x^2 - 2h\rho S_y S_x) + \left(L - \frac{1}{n} \right) (k^2 S_x^2 - 2k\rho S_y S_x) \quad \dots(4.3)$$

Proof: $V(T_{v7}) = E[T_{v7} - \bar{Y}]^2 = E[\bar{Y}e_1 + h\bar{X}(e_3' - e_3) + k\bar{X}(e_3 - e_2)]^2$
 $= E[\bar{Y}^2 e_1^2 + h^2 \bar{X}^2 (e_3' - e_3)^2 + k^2 \bar{X}^2 (e_3 - e_2)^2 + 2h\bar{Y}\bar{X}(e_3' - e_3)e_1$
 $+ 2hk\bar{X}^2 (e_3' - e_3)(e_3 - e_2) + 2k\bar{Y}\bar{X}(e_3 - e_2)e_1]$
 $= E[\bar{Y}^2 e_1^2 + h^2 \bar{X}^2 (e_3'^2 + e_3^2 - 2e_3 e_3') + k^2 \bar{X}^2 (e_3^2 + e_2^2 - 2e_2 e_3)$
 $+ 2h\bar{Y}\bar{X}(e_1 e_3' - e_1 e_3) + 2hk \bar{X}^2 (e_3 e_3' - e_3^2 - e_2 e_3' + e_2 e_3) + 2k\bar{Y}\bar{X}(e_1 e_3 - e_1 e_2)]$

$$= \left(L - \frac{1}{n'}\right) S_y^2 + \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N}\right) (h^2 S_x^2 - 2h\rho S_y S_x) + \left(L - \frac{1}{n}\right) (k^2 S_x^2 - 2k\rho S_y S_x)$$

(4) The minimum variance of the T_{v7} is

$$[V(T_{v7})]_{Min} = \left[\left(L - \frac{1}{n'}\right) - \left(L - \frac{2}{n'} + \frac{1}{N}\right) \rho^2 \right] S_y^2 \quad \dots(4.4)$$

Proof: By differentiating (4.3) with respect to h and k then equate to zero

$$\frac{d}{dh} [V(T_{v7})] = 0 \Rightarrow h = \rho \frac{S_y}{S_x}$$

and $\frac{d}{dk} [V(T_{v7})] = 0 \Rightarrow k = \rho \frac{S_y}{S_x}$

After replacing value of h and k in (4.3), we obtained

$$[V(T_{v7})]_{Min} = \left[\left(L - \frac{1}{n'}\right) - \left(L - \frac{2}{n'} + \frac{1}{N}\right) \rho^2 \right] S_y^2$$

Theorem 4.2:

(5) The estimator T_{v8} in terms of e_1, e_2, e_3 and e_3' is :

$$T_{v8} = \bar{Y} [1 + e_1 + \theta(e_3 - e_2 - e_1 e_2 + e_1 e_3 + (1 - 2\theta)e_2 e_3 + \theta e_2^2 - (1 - \theta)e_3^2)] \quad \dots(4.5)$$

Proof:

$$T_{v8} = \frac{\bar{y}_1 \bar{x}}{\theta \bar{x}_1 + (1 - \theta) \bar{x}_n} = \frac{\bar{Y} \bar{X} (1 + e_1) (1 + e_3)}{\theta \bar{X} (1 + e_2) + (1 - \theta) \bar{X} (1 + e_3)}$$

$$= \bar{Y} (1 + e_1) (1 + e_3) (1 + e_3 + \theta e_2 - \theta e_3)^{-1}$$

$$= \bar{Y} (1 + e_1) (1 + e_3) (1 + \theta e_2 + (1 - \theta) e_3)^{-1}$$

$$= \bar{Y} (1 + e_1) (1 + e_3) [1 - \theta e_2 - (1 - \theta) e_3 + \{\theta e_2 + (1 - \theta) e_3\}^2 - \dots]$$

$$= \bar{Y} [1 + e_1 + \theta(e_3 - e_2 - e_1 e_2 + e_1 e_3 + (1 - 2\theta)e_2 e_3 + \theta e_2^2 - (1 - \theta)e_3^2)]$$

(6) The bias of the estimator T_{v8} is

$$B(T_{v8}) = \bar{Y} \left(L - \frac{1}{n}\right) (\theta^2 C_x^2 - \theta \rho C_y C_x) \quad \dots(4.6)$$

Proof:

$$B(T_{v8}) = E[T_{v8} - \bar{Y}] = \bar{Y} E[1 + e_1 + \theta(e_3 - e_2 - e_1 e_2 + e_1 e_3 + (1 - 2\theta)e_2 e_3 + \theta e_2^2 - (1 - \theta)e_3^2) - 1]$$

$$= \bar{Y} \left(L - \frac{1}{n}\right) (\theta^2 C_x^2 - \theta \rho C_y C_x)$$

(7) Mean squared error of T_{v8} is

$$M(T_{v8}) = \bar{Y}^2 \left[\left(L - \frac{1}{n'}\right) C_y^2 + \left(L - \frac{1}{n}\right) (\theta^2 C_x^2 - 2\theta \rho C_y C_x) \right] \quad \dots(4.7)$$

Proof:

$$M(T_{v8}) = E[T_{v8} - \bar{Y}]^2 = \bar{Y}^2 E[1 + e_1 + \theta(e_3 - e_2 - e_1 e_2 + e_1 e_3 + (1 - 2\theta)e_2 e_3 + \theta e_2^2 - (1 - \theta)e_3^2) - 1]^2$$

$$= \bar{Y}^2 E[e_1^2 + \theta^2 (e_3^2 + e_2^2 - 2e_2 e_3) + 2\theta (e_1 e_3 - e_1 e_2)]$$

$$= \bar{Y}^2 \left[\left(L - \frac{1}{n'}\right) C_y^2 + \left(L - \frac{1}{n}\right) (\theta^2 C_x^2 - 2\theta \rho C_y C_x) \right]$$

(8) The minimum m.s.e. of T_{v8} is

$$[M(T_{v8})]_{Min} = \left[\left(L - \frac{1}{n'}\right) - \left(L - \frac{1}{n}\right) \rho^2 \right] S_y^2 \text{ when } \theta = \rho \frac{C_y}{C_x} \quad \dots(4.8)$$

Proof: By differentiating (4.7) with respect to θ then equate to zero

$$\frac{d}{d\theta} [M(T_{v8})] = 0 \Rightarrow \theta = \rho \frac{C_y}{C_x}$$

After replacing value of θ in (4.7), we obtained

$$[M(T_{v8})]_{Min} = \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{1}{n} \right) \rho^2 \right] S_y^2$$

Theorem 4.3:

(9) The estimator T_{v9} in terms of e_1, e_2, e_3 and e_3' is :

$$T_{v9} = \bar{Y} \left[1 + e_1 + \phi(e_3' - e_3 + e_1 e_3' - e_1 e_3) - (1 + 2\phi)e_3 e_3' + \phi e_3^2 + (1 + \phi)e_3'^2 \right] \quad \dots(4.9)$$

Proof: $T_{v9} = \frac{\bar{y}_1 \bar{x}'}{\phi \bar{x}_n + (1 - \phi) \bar{x}} = \frac{\bar{Y} \bar{X} (1 + e_1) (1 + e_3')}{\phi \bar{X} (1 + e_3) + (1 - \phi) \bar{X} (1 + e_3')} = \bar{Y} (1 + e_1) (1 + e_3') (1 + e_3 + \phi e_3 - \phi e_3')^{-1}$
 $= \bar{Y} (1 + e_1) (1 + e_3') (1 + \phi e_3 + (1 - \phi) e_3')^{-1} = \bar{Y} (1 + e_1) (1 + e_3') (1 - \phi e_3 - (1 - \phi) e_3' + \{\phi e_3 + (1 - \phi) e_3'\}^2 - \dots)^{-1}$
 $= \bar{Y} \left[1 + e_1 + \phi(e_3' - e_3 + e_1 e_3' - e_1 e_3) - (1 + 2\phi)e_3 e_3' + \phi e_3^2 + (1 + \phi)e_3'^2 \right]$

(10) The bias of the estimator T_{v9} is

$$B(T_{v9}) = \bar{Y} \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) (\phi^2 C_x^2 - \phi_2 \rho C_y C_x) \quad \dots(4.10)$$

Proof: $B(T_{v9}) = E[T_{v9} - \bar{Y}]$
 $= \bar{Y} E \left[1 + e_1 + \phi(e_3' - e_3 + e_1 e_3' - e_1 e_3) - (1 + 2\phi)e_3 e_3' + \phi e_3^2 + (1 + \phi)e_3'^2 - 1 \right]$
 $= \bar{Y} \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) (\phi^2 C_x^2 - \phi_2 \rho C_y C_x)$

(11) Mean squared error of T_{v9} is

$$M(T_{v9}) = \bar{Y}^2 \left[\left(L - \frac{1}{n'} \right) C_y^2 + \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) (\phi^2 C_x^2 - 2\phi_2 \rho C_y C_x) \right] \quad \dots(4.11)$$

Proof: $M(T_{v9}) = E[T_{v9} - \bar{Y}]^2$
 $= \bar{Y}^2 E \left[1 + e_1 + \phi(e_3' - e_3 + e_1 e_3' - e_1 e_3) - (1 + 2\phi)e_3 e_3' + \phi e_3^2 + (1 + \phi)e_3'^2 - 1 \right]^2$
 $= \bar{Y}^2 E \left[e_1^2 + \phi^2 (e_3'^2 + e_3^2 - 2e_3 e_3') + 2\phi_2 (e_1 e_3' - e_1 e_3) \right]$
 $= \bar{Y}^2 \left[\left(L - \frac{1}{n'} \right) C_y^2 + \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) (\phi^2 C_x^2 - 2\phi_2 \rho C_y C_x) \right]$

(12) The minimum m.s.e. of T_{v9} is

$$[M(T_{v9})]_{Min} = \left[L - \left(\frac{1}{n'} \right) - \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2 \text{ when } \phi = \rho \frac{C_y}{C_x} \quad \dots(4.12)$$

Proof: By differentiating (4.11) with respect to ϕ then equate to zero

$$\frac{d}{d\phi} [M(T_{v9})] = 0 \Rightarrow \phi = \rho \frac{C_y}{C_x}$$

After replacing value of ϕ in (4.11), we obtained

$$[M(T_{v9})]_{Min} = \left[\left(L - \frac{1}{n'} \right) - \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2$$

Theorem 4.4:

(13) The estimator T_{v10} in terms of e_1, e_2, e_3 and e_3' is :

$$T_{v10} = \bar{Y} \left[1 + e_1 + \psi(e_3' - e_2 + e_1 e_3' - e_1 e_2 + \psi e_2^2 + \psi e_3'^2 - e_3'^2 - e_2 e_3') \right] \quad \dots(4.13)$$

Proof: $T_{v10} = \frac{\bar{y}_1 \bar{x}'}{\psi \bar{x}_1 + (1 - \psi) \bar{x}} = \frac{\bar{Y} \bar{X} (1 + e_1) (1 + e_3')}{\psi \bar{X} (1 + e_2) + (1 - \psi) \bar{X} (1 + e_3')} = \bar{Y} (1 + e_1) (1 + e_3') (1 + e_3' + \psi e_2 - \psi e_3')^{-1}$

$$= \bar{Y}(1 + e_1)(1 + e_3)(1 + \psi e_2 + (1 - \psi)e_3)^{-1} = \bar{Y}(1 + e_1)(1 + e_3)(1 - \psi e_2 - (1 - \psi)e_3 + \{\psi e_2 + (1 - \psi)e_3\}^2 - \dots)$$

$$= \bar{Y}[1 + e_1 + \psi(e_3 - e_2 + e_1 e_3 - e_1 e_2 + \psi e_2^2 + \theta_3 e_3^2 - e_3^2 - e_2 e_3)]$$

(14) The bias of the estimator $T_{v_{10}}$ is

$$B(T_{v_{10}}) = \bar{Y} \left(\psi^2 \left(L - \frac{1}{N} \right) C_x^2 - 2\psi \left(\frac{1}{n'} - \frac{1}{N} \right) C_x^2 - \psi \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho C_y C_x \right) \quad \dots(4.14)$$

Proof: $B(T_{v_{10}}) = E[T_{v_{10}} - \bar{Y}] = \bar{Y} E[1 + e_1 + \psi(e_3 - e_2 + e_1 e_3 - e_1 e_2 + \psi e_2^2 + \psi e_3^2 - e_3^2 - e_2 e_3) - 1]$

$$= \bar{Y} \left(\psi^2 \left(L - \frac{1}{N} \right) C_x^2 - 2\psi \left(\frac{1}{n'} - \frac{1}{N} \right) C_x^2 - \psi \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho C_y C_x \right)$$

(15) Mean squared error of $T_{v_{10}}$ is

$$M(T_{v_{10}}) = \bar{Y}^2 \left[\left(L - \frac{1}{n'} \right) C_y^2 + \left(L - \frac{2}{n'} + \frac{1}{N} \right) (\psi^2 C_x^2 - 2\psi \rho C_y C_x) \right] \quad \dots(4.15)$$

Proof: $M(T_{v_{10}}) = E[T_{v_{10}} - \bar{Y}]^2$

$$= \bar{Y}^2 E[1 + e_1 + \psi(e_3 - e_2 + e_1 e_3 - e_1 e_2 + \psi e_2^2 + \psi e_3^2 - e_3^2 - e_2 e_3) - 1]^2$$

$$= \bar{Y}^2 E[e_1^2 + \psi^2(e_3^2 + e_2^2 - 2e_2 e_3) + 2\psi(e_1 e_3 - e_1 e_2)]$$

$$= \bar{Y}^2 \left[\left(L - \frac{1}{n'} \right) C_y^2 + \left(L - \frac{2}{n'} + \frac{1}{N} \right) (\psi^2 C_x^2 - 2\psi \rho C_y C_x) \right]$$

(16) The minimum m.s.e. of $T_{v_{10}}$ is

$$[M(T_{v_{10}})]_{Min} = \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2 \text{ when } \psi = \rho \frac{C_y}{C_x} \quad \dots(4.16)$$

Proof: By differentiating (4.15) with respect to ψ then equate to zero

$$\frac{d}{d\psi} [M(T_{v_{10}})] = 0 \Rightarrow \psi = \rho \frac{C_y}{C_x}$$

After replacing value of ψ in (4.15), we obtained

$$[M(T_{v_{10}})]_{Min} = \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2$$

V. Comparisons

In this section we derived the conditions under which the suggested estimators are superior to the Ahmed et al. (2006).

(1) $D_7 = \min [M(t_7)] - \min [M(T_{v_7})]$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(\frac{1}{n_1} - \frac{1}{N} \right) \rho^2 \right] S_y^2 - \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2$$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(L - \frac{1}{n'} \right) \right] S_y^2 + \left[- \left(\frac{1}{n_1} - \frac{1}{N} \right) + \left(L - \frac{2}{n'} - \frac{1}{N} \right) \right] \rho^2 S_y^2$$

$$= \left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right] S_y^2 - \left[\frac{1}{n_1} - \frac{2}{N} - L + \frac{2}{n'} \right] \rho^2 S_y^2$$

(T_{v_7}) is better than t_7 , if

$$D_7 > 0 \Rightarrow \rho^2 < \frac{\left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right]}{\left[\frac{1}{n_1} - \frac{2}{N} - L + \frac{2}{n'} \right]} \Rightarrow \rho^2 < 1 + \frac{nn_1^2(N-n')(N-1)}{n'N(N-n)(n-n_1) - 2nn_1^2(N-n')(N-1)}$$

$$\Rightarrow \rho < \pm \sqrt{1+Q} \quad \Rightarrow -\sqrt{1+Q} < \rho < +\sqrt{1+Q}$$

where $Q > 1 \Rightarrow nn_1^2(N-n')(N-1) > n'N(N-n)(n-n_1) - 2nn_1^2(N-n')(N-1)$

(2) $D_8 = \min[M(t_8)] - \min[M(T_{v8})]$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(\frac{1}{n_1} - \frac{1}{n} \right) \rho^2 \right] S_y^2 - \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{1}{n} \right) \rho^2 \right] S_y^2 = \left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right] S_y^2 - \left[\frac{1}{n_1} - L \right] \rho^2 S_y^2$$

(T_{v8}) is better than t_8 , if

$$D_8 > 0 \quad \Rightarrow \left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right] S_y^2 - \left[\frac{1}{n_1} - L \right] \rho^2 S_y^2 > 0$$

$$\Rightarrow \left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right] S_y^2 - \left[\frac{1}{n_1} - L \right] \rho^2 S_y^2 > 0 \Rightarrow \rho^2 < \frac{\left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right]}{\left[\frac{1}{n_1} - L \right]}$$

$$\Rightarrow \rho^2 < \left[\frac{(N-n)(n-n_1)n'N}{nn_1^2(N-1)(N-n')} \right]^{-1} - 1 \Rightarrow \rho < \pm \sqrt{\frac{1-M}{M}}$$

where $M < 1 \Rightarrow (N-n)(n-n_1)n'N < nn_1^2(N-1)(N-n')$

(3) $D_9 = \min[M(t_9)] - \min[M(T_{v9})]$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(\frac{1}{n} - \frac{1}{N} \right) \rho^2 \right] S_y^2 - \left[\left(L - \frac{1}{n'} \right) - \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2$$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(L - \frac{1}{n'} \right) \right] S_y^2 + \left[-\left(\frac{1}{n} - \frac{1}{N} \right) + \left(\frac{1}{n} - \frac{2}{n'} + \frac{1}{N} \right) \right] \rho^2 S_y^2$$

$$= \left[\frac{1}{n_1} - \frac{1}{N} - L - \frac{1}{n'} \right] S_y^2 - 2 \left[\frac{1}{n} - \frac{1}{N} \right] \rho^2 S_y^2$$

(T_{v9}) is better than t_9 , if

$$D_9 > 0 \Rightarrow \left[\frac{1}{n_1} - \frac{1}{N} - L - \frac{1}{n'} \right] S_y^2 - 2 \left[\frac{1}{n} - \frac{1}{N} \right] \rho^2 S_y^2 > 0 \Rightarrow \rho^2 < \frac{1}{2} \frac{\left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right]}{\left[\frac{1}{n} - \frac{1}{N} \right]}$$

$$\Rightarrow \rho^2 < \frac{1}{2} - \frac{(N-n)(n-n_1)n'N}{nn_1^2(N-1)(N-n')} \Rightarrow \rho < \pm \sqrt{\frac{1}{2} - M} \Rightarrow -\sqrt{\frac{1}{2} - M} < \rho < +\sqrt{\frac{1}{2} - M}$$

where $M < \frac{1}{2} \Rightarrow 2(N-n)(n-n_1)n'N < nn_1^2(N-1)(N-n')$

(4) $D_{10} = \min[M(t_{10})] - \min[M(T_{v10})]$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(\frac{1}{n_1} - \frac{1}{N} \right) \rho^2 \right] S_y^2 - \left[\left(L - \frac{1}{n'} \right) - \left(L - \frac{2}{n'} + \frac{1}{N} \right) \rho^2 \right] S_y^2$$

$$= \left[\left(\frac{1}{n_1} - \frac{1}{N} \right) - \left(L - \frac{1}{n'} \right) \right] S_y^2 + \left[-\left(\frac{1}{n_1} - \frac{1}{N} \right) + \left(L - \frac{2}{n'} - \frac{1}{N} \right) \right] \rho^2 S_y^2$$

$$= \left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right] S_y^2 - \left[\frac{1}{n_1} - \frac{2}{N} - L + \frac{2}{n'} \right] \rho^2 S_y^2$$

(T_{v10}) is better than t_{10} , if $D_{10} > 0$

$$\Rightarrow \rho^2 < \frac{\left[\frac{1}{n_1} - \frac{1}{N} - L + \frac{1}{n'} \right]}{\left[\frac{1}{n_1} - \frac{2}{N} - L + \frac{2}{n'} \right]} \Rightarrow \rho^2 < 1 + \frac{nn_1^2(N-n')(N-1)}{n'N(N-n)(n-n_1) - 2nn_1^2(N-n')(N-1)}$$

$$\Rightarrow \rho < \pm \sqrt{1+Q} \Rightarrow -\sqrt{1+Q} < \rho < +\sqrt{1+Q}$$

where $Q > 1 \Rightarrow nn_1^2(N-n')(N-1) > n'N(N-n)(n-n_1) - 2nn_1^2(N-n')(N-1)$

VI. Numerical Illustrations

We consider two populations A and B, first one is the artificial population of size $N = 200$ [source Shukla et al. (2009a)] and another one is from Ahmed et al. (2006) with the following parameters:

Table 6.1 Parameters of Populations A and B

Population	N	\bar{Y}	\bar{X}	S_y^2	S_x^2	ρ	C_x	C_y
A	200	42.485	18.515	199.0598	48.5375	0.8652	0.3763	0.3321
B	8306	253.75	343.316	338006	862017	0.522231	2.70436	2.29116

Let $n = 60$, $n = 40$, $n_1 = 35$ for population A and $n = 2000$, $n = 500$, $n_1 = 450$ for population B respectively. Then the bias and M.S.E of suggested estimators (using the expressions of bias and m.s.e. of Section 5) and other existing estimators with Ahmed et al. (2006) methods are given in table 6.2 and 6.3 for population A and B respectively.

Table 6.2 Bias and MSE for Population A and B

Estimators	Population A		Population B	
	Bias	MSE	Bias	MSE
T_{v7}	0	2.338387	0	458.4694
T_{v8}	-0.000001	1.841686	0.000003	561.7505
T_{v9}	0.000001	2.882792	0.000001	478.9972
T_{v10}	-0.025350	2.338387	-0.347570	458.4694

Table 6.3 Bias and MSE for Population A and B for Ahmed et al. (2006)

Estimators	Population A		Population B	
	Bias	MSE	Bias	MSE
\bar{y}_r	0	4.692124	0	710.4302
\bar{y}_{RAT}	0.005080	4.908211	0.22994	768.7752
\bar{y}_{COMP}	0.003879	4.188044	0.050411	689.9429
t_7	0	1.179736	0	516.6780
t_8	-0.000001	4.159944	0.000003	689.9452
t_9	-0.000006	1.711916	0.000002	537.1631
t_{10}	-0.000008	1.179736	0.000003	516.6780

The sampling efficiency of suggested estimators over Ahmed et al. (2006) is defined as:

$$E_i = \frac{Opt[M(T_{vi})]}{Opt[M(t_i)]}; \quad i = 7,8,9,10 \quad \dots(6.1)$$

The efficiency for population A and population B are given in table 6.4

Table 6.4 Efficiency for Population A and B over Ahmed et al. (2006)

Efficiency	Population A	Population B
E_7	1.982128	0.887341
E_8	0.442719	0.814196
E_9	1.683957	0.891717
E_{10}	1.982128	0.887341

VII. Discussion

The idea of two-phase sampling is used while considering, the auxiliary population mean is unknown and numbers of available observations are considered as random variable. Some strategies are suggested in Section 3 and the estimator of population mean derived. Properties of derived estimators like bias and m.s.e are discussed in the Section 4. The optimum value of parameters of suggested estimators is obtained as well in same section. Ahmed et al. (2006) estimators are considered for comparison purpose and two populations **A** and **B** considered for numerical study first one from Shukla et al. (2009) and another one is Ahmed et al. (2006). The sampling efficiency of suggested estimator over Ahmed et al. (2006) is obtained and suggested strategy is found very close with Ahmed et al. (2006) when \bar{X} is not known.

VIII. Conclusions

The proposed estimators are useful when some observations are missing in the sampling and population mean of auxiliary information is unknown. For population **A** proposed estimators T_{v8} are found to be more efficient than the existing estimators. For population **B** proposed estimators T_{v7}, T_{v8}, T_{v9} and T_{v10} are found to be more efficient than the existing estimators.

REFERENCES

- [1] Ahmed, M.S., Al-Titi, O., Al-Rawi, Z. and Abu-Dayyeh, W. (2006): *Estimation of a population mean using different imputation methods*, Statistics in Transition, 7, 6, 1247-1264.
- [2] Hansen, M. H. and Hurwitz, W. N. (1946): *The problem of non-response in Sample Surveys*. Journal of the American Statistical Association, 41, 236, 517-529.
- [3] Heitjan, D. F. and Basu, S. (1996): *Distinguishing 'Missing at random' and 'Missing completely at random'*. The American Statistician, 50, 207-213.
- [4] Kalton, G., Kasprzyk, D. and Santos, R. (1981): *Issues of non-response and imputation in the Survey of Income and Program Participation*. Current Topics in Survey Sampling, (D. Krewski, R. Platek and J.N.K. Rao, eds.), 455-480, Academic Press, New York.
- [5] Little, R.J.A., and Rubin, D.B. (1987): *Statistical Analysis with Missing Data*, John Wiley and Sons, NY.
- [6] Lee, H., Rancourt, E. and Sarndal, C. E. (1994): *Experiments with variance estimation from survey data with imputed values*. Journal of official Statistics, 10, 3, 231-243.
- [7] Lee, H., Rancourt, E. and Sarndal, C. E. (1995): *Variance estimation in the presence of imputed data for the generalized estimation system*. Proc. of the American Statist. Assoc. (Social Survey Research Methods Section), 384-389.
- [8] Rao, J. N. K. and Sitter, R. R. (1995): *Variance estimation under two-phase sampling with application to imputation for missing data*, Biometrika, 82, 453-460.
- [9] Rubin, D. B. (1976): *Inference and missing data*. Biometrika, 63, 581-593.
- [10] Sande, I. G. (1979): *A personal view of hot deck approach to automatic edit and imputation*. Journal Imputation Procedures. Survey Methodology, 5, 238-246.
- [11] Shukla, D. (2002): *F-T estimator under two-phase sampling*, Metron, 59, 1-2, 253-263.
- [12] Shukla, D. and Dubey, J. (2001): *Estimation in mail surveys under PSNR sampling scheme*. Ind. Soc. Ag. Stat., 54, 3, 288-302.
- [13] Shukla, D. and Dubey, J. (2004): *On estimation under post-stratified two phase non-response sampling scheme in mail survey*. International Journal of Management and Systems, 20, 3, 269-278.
- [14] Shukla, D. and Dubey, J. (2006): *On earmarked strata in post-stratification*. Statistics in Transition, 7, 5, 1067-1085.
- [15] Shukla, D. and Thakur, N. S. (2008): *Estimation of mean with imputation of missing data using factor-type estimator*, Statistics in Transition, 9, 1, 33-48.
- [16] Shukla, D., Thakur, N. S., Pathak, S. and Rajput D. S. (2009): *Estimation of mean with imputation of missing data using factor-type estimator in two-phase sampling*, Statistics in Transition, 10, 3, 397-414.
- [17] Shukla, D., Thakur, N. S., Thakur, D. S. (2009a): *Utilization of non-response auxiliary population mean in imputation for missing observations*, Journal of Reliability and Statistical Studies, 2, 28-40.
- [18] Shukla, D., Thakur, N. S., Thakur, D. S. and Pathak, S. (2011): *Linear combination based imputation method for missing data in sample*, International Journal of Modern Engineering Research, 1, 2, 580-596.
- [19] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Ashok, C. (1984): *Sampling Theory of Surveys with Applications*, Iowa State University Press, I.S.A.S. Publication, New Delhi.
- [20] Singh, S. and Horn, S. (2000): *Compromised imputation in survey sampling*, Metrika, 51, 266-276.
- [21] Thakur, N. S., Yadav, K. and Pathak, S. (2011): *Estimation of mean in presence of missing data under two-phase sampling scheme*, Journal of Reliability and Statistical Studies, 4, 2, 93-104.
- [22] Thakur, N. S., Yadav, K. and Pathak, S. (2012): *Some imputation methods in double sampling scheme for estimation of population mean*, International Journal of Modern Engineering Research, 2, 1, pp-200-207.

Handwritten Characters Classification Using Neural Networks and Moments Features

Mohamed Abaynarh¹, Hakim Elfadili² and Lahbib Zenkour¹

¹(Ecole Mohammadia d'Ingenieurs, Avenue Ibsina B.P. 765 Agdal Rabat, Morocco)

²(Ecole Nationale des Sciences Appliquées, Quartier Industriel Ain Chkef Route Bensouda - Fès BP 72 Fès Principale, Morocco)

ABSTRACT: This paper presents unconstrained handwritten Amazighe character recognition based upon orthogonal moments and neural networks classifier. The proposed system extracts moments features from character images. A total of 6600 images were considered for experimentation and overall accuracy found to be 97.46%. The novelty of the proposed method is independence of size, slant, orientation, and translation. The performance of the proposed method is experimentally evaluated and the promising results and findings are presented. Our method is compared to an Euclidian distance criterion and K-NN classifier algorithm, results show performances of our method.

Keywords: Neural Network, character recognition, orthogonal Moments

I. Introduction

Character recognition is one of the most challenging topics in pattern recognition, In the past several decades, a large number of OCR systems have been developed for natural languages [1-3]. Handwritten character recognition is difficult due to the large category set, wide variability of writing styles, and the confusion between similar characters.

Artificial neural networks trained with back propagation algorithm are frequently used in the field of pattern recognition and have shown its power and good performance for pattern recognition. However, the performance of those classifiers is strongly affected by the quality of the representation of the pattern i.e. features. Consequently, we present an efficient feature extraction method based on orthogonal Legendre moments.

Hu [4] first introduced the use of seven invariants moments which are defined on moments of the image as features for pattern classification. These moments are nonlinear and invariant under translation, rotation, scaling, and image reversal. Where lower orders of moments are not enough to classify patterns, higher orders will be used, although, the higher orders resulted in higher sensitivity.

In this paper, we use the Maximum entropy principle MEP as feature selection criterion which produces finite optimal moment orders carrying out only moments of low orders containing sufficient and pertinent information needed for classification.

Our classification method is used with different network topologies, compared to other methods as minimum mean distance, nearest neighbor.

In general, the overall recognition process can be divided into 3 main sections, namely segmentation, preprocessing, and classification [6]. Segmentation requires isolating the characters individually before they are fed to the

preprocessing unit where the important features of characters (feature extraction) are identified. Finally, classification process is done by determining the category or the group of each character used during the recognition process.

Experimental results show that the proposed method reduces the computational burden of the recognition system in terms of the total number of layers and nodes, while showing improved performances in terms of recognition rate and generalization ability

The rest of this paper is organized as follows: In the coming section, we describe briefly the database used in our system. Section III points out the proposed method of moments features extraction. Section IV explains neural classifier. Section V is devoted to experimental results. Finally, Section VI draws conclusion and summarizes the paper.

II. Database preparation

2.1. Amazighe language

The Amazighe (Berber) language is spoken in Morocco, Algeria, Tunisia, Libya, and Siwa (an Egyptian Oasis); it is also spoken by many other communities in parts of Niger, Burkina Faso and Mali [5]. It is used by tens of millions of people in North Africa mainly for oral communication and has been integrated in mass media and in

As far as the alphabet is concerned, and because of historical and cultural reasons, Tifinaghe has become the official graphic system for writing Amazighe [6]. IRCAM kept only pertinent phonemes for Tamazighte, so the number of the alphabetical phonetic entities is 33, but Unicode codes only 31 letters plus a modifier letter to form the two phonetic units: $\square \square(g^w)$ and $\square \square(k^w)$. [7]. The Fig.1 represents the repertoire of Tifinaghe which is recognized and used in Morocco with their correspondents in Latin characters.

o	ⵝ	ⵉ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ
ya	yab	yag	yag ^w	yad	yad	yey	yaf	yak	yak ^w	yah
a	b	g	g ^w	d	d	e	f	k	k ^w	h
[a]	[b/β]	[g/ǧ]	[g ^w]	[d/ð]	[d]	[e]	[f]	[k/ç]	[k ^w]	[h]
ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ
yah	yac	yax	yaq	yi	yaj	yal	yam	yan	yu	yar
h	x	q	i	j	l	m	n	u	r	
[h]	[ç]	[x]	[q]	[i]	[j]	[l]	[m]	[n]	[u]	[r]
ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ	ⵏ
yar	yagh	yas	yas ^w	yac	yat	yat	yaw	yay	yaz	yaz
r	gh	s	s ^w	c	t	t	w	y	z	z
[r]	[ʃ]	[s]	[s ^w]	[j]	[t]	[t]	[w]	[j]	[z]	[z]

Figure 1: Neo-Tifinaghe alphabet as used in Morocco with their Correspondents in Latin Characters.

2.2. Data collection

Our database will facilitate fruitful research on handwritten recognition of Amazighe through free access to the researchers. Descriptions of these components of the present database are given below.

The database contains forms of unconstrained handwritten characters including 7524 isolated characters, gathered from 57 different and independent writers, The whole set of available isolated characters datas have been split into a training set consisting of 6600 samples and a test set consisting of 924 samples. Before collection of datas, the following points were decided to make the database as much representative as possible. Common factors responsible for variations in handwriting styles include age, sex, education, profession, writing instrument, writing surface. No restriction was imposed on the writers except for specifying rectangular regions for writing isolated characters of different sizes. Since such rectangular regions are large enough, the restriction may not be considered as a serious one.

The forms are scanned at 300 d.p.i. and stored as grayscale BMP images of 100 × 100 size. Samples of isolated characters from the present database are shown in Fig. 2. In some cases a character may touches or crosses the horizontal or vertical lines of the bounding box. Therefore two types of errors may happen. In the major error case, some character’s dots or some complementary strokes of it were omitted and the result was not distinguishable, but in the minor error case, usually the last part of the character was missed.

The characters prepared as explained in previous Section, are scanned using a scanner and these characters will be segregated according to their own character group. One example is shown below in Fig. 2.

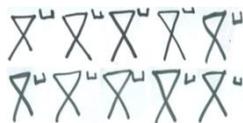


Figure 2: Sample of character

Note that the scanned images are in RGB scale.

These images have to be converted into grayscale format before further processing can be done. Using appropriate grayscale thresholding, binary images are to be created.

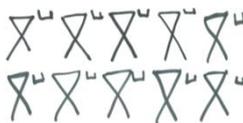


Figure 3: Binary image of sample character

Fig. 3 shows the generated binary image using image processing tool box in MATLAB and 8-connectivity analysis.

III. Features extraction

For extracting the feature, the moment based approach is proposed. The most important aspect of handwritten recognition scheme is the selection of good feature set,

which is reasonably invariant with respect to shape variations caused by various writing styles. The major advantage of this approach stems from its robustness to small variation, ease of implementation and provides good recognition rate. Moments based feature extraction method provides good result even when certain preprocessing steps like filtering, smoothing and slant removing are not considered. Especially, the advantages of considering orthogonal moments are that they are shift, and scale invariants and are very robust in the presence of noise [8-9]. The invariant properties of moments are utilized as pattern sensitive features in classification and recognition applications [10,11].

In this section, we explain the concept of feature extraction method used for extracting features for efficient classification and recognition. The following paragraph explains in detail about the feature extraction methodology. Statistical moments represent average values of processes (powered to order *n*) when a random variable is involved. Here, the original images were considered as two dimensional arrays of a random variable of dimension *N*×*N*. The random variables took values from level 0 to 255, as the images were considered in gray levels quantized in 8 bytes

(Gray levels were obtained from BMP format). Moments were calculated for the random variable *X*, which was identified with the image block. In addition, *X* is a matrix of two coordinates (*x*, *y*) obtained from the image matrix *f*(*x*, *y*). The definition of (*p*+*q*) order invariant moment around the origin is given by:

The Legendre moments of order (*p* + *q*) are defined for a given real image intensity function *f*(*x*, *y*) as

$$\lambda_{p,q} = \frac{(2p+1)(2q+1)}{4} \int_R \int_R P_p(x)P_q(y)f(x,y) dx dy \quad (1)$$

Where *f*(*x*, *y*) is assumed to have bounded support The Legendre polynomials *P_p*(*x*) are a complete orthogonal basis set on the interval [−1,1] for an order *p* they are defined as

$$p_p(x) = \frac{1}{2^p p!} \frac{d^p}{dx^p} (x^2 - 1)^p \quad (2)$$

The orthogonality property is guaranteed by the equality:

$$\int_{-1}^1 p_p(x)p_q(x)dx = \frac{2}{(2p+1)} \delta_{p,q} \quad (3)$$

Where $\delta_{p,q}$ is the Kronecker function, that is,

$$\delta_{p,q} = \begin{cases} 1 & \text{if } p = q \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

4-1-Image reconstruction by Legendre moments

By taking the orthogonality principle into consideration, the image function *f*(*x*, *y*) can be written as an infinite series expansion in terms of Legendre polynomials over the square [−1,1]×[−1,1]:

$$f(x,y) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \lambda_{p,q} P_p(x)P_q(y), \quad (5)$$

Where the Legendre moments are computed over the same square

If only Legendre moments of order smaller than or equal to θ are given, then the function $f(x, y)$ can be approximated by a continuous function which is a truncated series:

$$f_{\theta}(x, y) = \sum_{p=0}^{\theta} \sum_{q=0}^p \lambda_{p,q} P_{p-q}(x) P_q(y), \quad (6)$$

Furthermore, $\lambda'_{p,q}$ must be replaced by their numerical approximation which will be pointed out on the following section. The number of moments used in the reconstruction of image for a given θ is defined by

$$N_{total} = \frac{(\theta+1)(\theta+2)}{2}, \quad (7)$$

3-1-Approximation of the Legendre moments

In practice the Legendre moments have to be computed from sampled data, that is, the rectangular sampling of the original image function $f(x, y)$, producing the set of samples $f(x_i, y_j)$ with an (M, N) array of pixels, thus we define the discrete version of $\lambda_{p,q}$ in terms of summation by the traditional commonly used formula :

$$\tilde{\lambda}_{p,q} = \frac{(2p+1)(2q+1)}{4} \sum_{i=1}^M \sum_{j=1}^N P_p(x_i) P_q(y_j) f(x_i, y_j) \Delta x \quad (8)$$

Where $\Delta x = (x_i - x_{i-1})$ and $\Delta y = (y_j - y_{j-1})$ are sampling intervals in the x and y directions.

It is clear, however, that $\tilde{\lambda}_{p,q}$ is not a very accurate approximation of $\lambda_{p,q}$, in particular, when the moment order $(p + q)$ increases

The piecewise constant approximation of $f(x, y)$ proposed recently by Liao and Pawlak [21, 37], yields the following approximation of $\lambda_{p,q}$:

$$\tilde{\lambda}_{p,q} = \sum_{i=1}^M \sum_{j=1}^N H_{p,q}(x_i, y_j) f(x_i, y_j), \quad (9)$$

With the supposition that $f(x, y)$ is piecewise constant over the interval

$$\left[x_i - \frac{\Delta x}{2}, x_i + \frac{\Delta x}{2} \right] \times \left[y_j - \frac{\Delta y}{2}, y_j + \frac{\Delta y}{2} \right] \quad (10)$$

And where

$$H_{p,q}(x_i, y_j) = \frac{(2p+1)(2q+1)}{4} \int_{x_i-\Delta x/2}^{x_i+\Delta x/2} \int_{y_j-\Delta y/2}^{y_j+\Delta y/2} P_p(x) P_q(y) dx dy, \quad (11)$$

represents the integration of the polynomial $P_p(x) P_q(y)$ around the (x_i, y_j) pixel.

This approximation allows a good quality of reconstructed images by reducing the reconstruction error.

In this paper, we determine the order of the truncated expansion of $f_{\theta}(x, y)$ which provides a good quality of the reconstructed object. The moments used in this reconstruction process will constitute the optimal subset for representing this object. Then, we introduce the Maximum Entropy Principle (MEP) to extract relevant moments that

uniquely represent the patterns [12], [13], [14]. By applying the Maximum Entropy Principle the entropy function monotonically increases up to a certain optimal order where sufficient image information is recreated and then become relatively constant [12].

IV. Classification and recognition

4.1. Minimum Mean Distance

Minimum mean distance is a conventional nonparametric statistical classifier. At first, it makes mean feature vector of each class using training samples. Then in the testing stage, it assigns an unknown input pattern to the class which has minimum distance with corresponding mean feature vector among the all classes. Different definitions of distance can be used practically e.g. Euclidian distance, city-block distance, and so on.

In the minimum distance classifier, each character class, C_k , is represented with the sample means, μ_k , learned from the training examples. When a new example is given, it is compared to each character class by calculating the Euclidean distance. The example is assigned to class k for which the distance is minimum.

The training example of class k , c_k , with the smallest distance to the test example, a , is the nearest neighbour of a , the equation is shown below:

$$d(a, c^k) = \sum_{i=1}^m (a_i - c_i^k)^2 \quad (12)$$

4.2. K-Nearest-Neighbor (KNN)

Nearest neighbor rule is also a conventional nonparametric statistical classifier. In training stage, it stores all training samples in a table. Then in testing stage, it assigns an unknown input pattern to which class has minimum distance to a training sample of that class. Just such as minimum mean distance classifier, different definitions of distance can also be used here.

We used Legendre moments based features for recognition of Amazighe characters. In this work, we have considered only isolated characters.

K-Nearest-Neighbor (KNN) classifier is an effective technique for classification problems in which the pattern classes exhibit a reasonably limited degree of variability. The K-NN classifier is based on the assumption that the classification of an instance is most similar to the classification of other instances that are nearby in the vector space. It works by calculating the distances between one input patterns with the training patterns. A K-Nearest-Neighbor classifier takes into account only the k nearest prototypes to the input pattern. Usually, the decision is determined by the majority of class values of the k neighbors.

In the K-Nearest neighbor classification, we compute the distance between features of the test sample and the feature of every training sample. The class of majority among the k-nearest training samples is based on the Euclidian minimum distance. The classification is carried using K-NN as follows:

Algorithm

Input: Isolated Binary Amazighe character Images

Output: Recognition of the Character

Method: Legendre Moments, and K-NN Classifier

1. Preprocess the input image to eliminate the noise using median filter.
2. Fit the bounding box on an input image and crop the image, then resize it to 100x100 pixels.
3. Extract the moment based features stored in a feature vector.
4. A Euclidian distance criterion and K-NN classifier used to classify the test sample
5. Stop

4.3. Multi Layer Perceptron

Artificial neural network (ANN) has been inspired from biological neural structure of human brain. Although ANN is a very simple abstraction of its biological counterpart, it has been interested a lot because of its extensive power in pattern classification and clustering in the recent years [15]. Multi layer perceptron is a feed-forward neural network with one or more layers of nodes between the input and output layers.

These in-between layers are called hidden layers. Each node in a layer is connected to the all nodes in the next layer. Using MLP in the context of a classifier requires all output nodes to be set to 0 except for the node that is marked to correspond to the class the input is from. That desired output is 1. MLP training is done using an iterative gradient descent procedure known as back-propagation algorithm [16].

Neural network is widely used as a classifier in many handwritten character recognition systems [10, 17]. Also, due to the simplicity, generality, and good learning ability of neural networks, these types of classifiers are found to be more efficient [10]. In this paper, multilayer feed forward neural network (MFNN) is used to classify the patterns. In our algorithm, the stochastic gradient algorithm as a minimization procedure is used during the learning phase. The weights are updated on the basis of a single sample.

The inputs of the MFNN are feature vectors derived from the proposed feature extraction method described in the previous section. The number of nodes in the output layer is set to the number of Amazighe characters classes.

Experiments were conducted using the initial weight vectors that have been randomly chosen from a uniform distribution in (-1, 1), this weight range has been used in [18, 19].

Structure of MLP network for English character recognition is shown in Figure 5.

In this paper, a neural network is used as a classifier in character recognition where the inputs to the neural network are feature vectors derived from the proposed feature extraction technique described in the previous section [20].

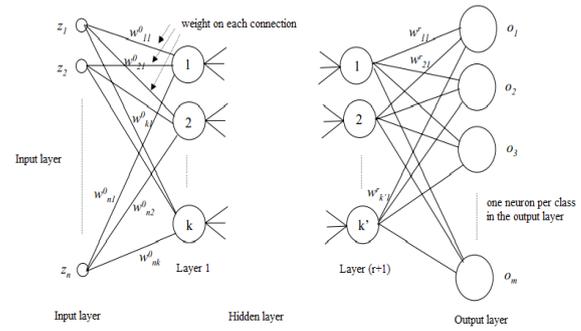


Figure 4: Multi-layer-Perceptron neuronal network

The output of each node is a pondered sum of its inputs:

$$o_i = \varphi(a_i) = \varphi(\sum_{k=1}^N (w_{ik} z_k)) \quad (13)$$

With Z_k the k^{th} component of sample vector. W_{ik} is the weight of the connection which rely unit k and unit i .

a_i is the activation of the unit i .

φ is the activation function of the units which is a threshold function with the following expression :

$$\varphi(x) = \begin{cases} -1, & x < \theta \\ +1, & x \geq \theta \end{cases} \quad (14)$$

Our procedure of handwritten Amazighe character recognition is given below

- Capture the scanned characters into 100x100 pixels
- Apply our proposed Feature Extraction method without any image preprocessing
- Implement the Neural Network Classifier with the subset already extracted
- Get the recognized character.

A complete flowchart of handwritten Amazighe character recognition is given below in Fig. 5.

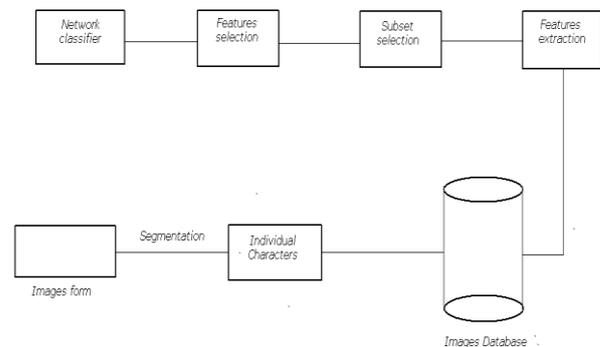


Figure 5: System for Amazighe character recognition

V. Experimental results

In this experiment, we are interested in determining how well the pre-trained recognizer works for a new user under classification methods. Each time, a different individual's data set is held out for a test set, and the neuronal classifier is trained with all other users' data and then tested on the holdout set. For each round, there are 6600 characters for training, and 924 characters for testing.

The training and testing data were different; even more the data used for testing were outside training set. The Training dataset consists of 200 samples for character, and the testing dataset consists of 28 samples for each character of different handwritten. Back propagation algorithm is used for the training of the Neural Network. At the training time, weight and bias will be updated in each iteration if there is a difference between the computed output and the target. Table 1 shows the recognition rate of multiples classifiers, the classifier recognition rate (%) is considered as the number of recognized characters in the training (test) phase over the total number of characters.

Table1: The results of error rate classification of different classifiers

Utilized classifier	Error rate classification(%)
Minimum Mean Distance(MMD)	83,5
Nearest Neighbor Rule(NN)	94,5
Artificial Neural Network(ANN)-Multi layer perceptron	97

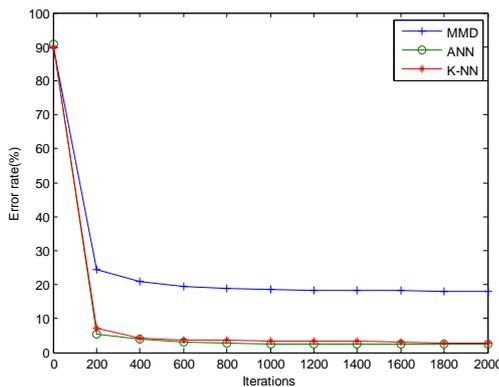


Figure 6: Error rate for the same set of characters of different classifiers: MMD, K-NN and ANN

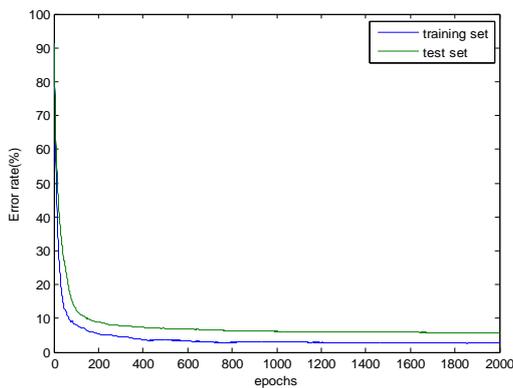


Fig 7. shows that our combined method of neural networks and moment features, gives the best results than the other methods in terms of recognition rate

Error rates on both sets usually go up and down simultaneously. However, if the neural network is trained again and again and more than the needed information is provided, the error rate on the training set continues to decrease but it will revert on the testing set. This situation is

called over-training. The relationship between error rate on training and testing sets is shown in Fig. 6.

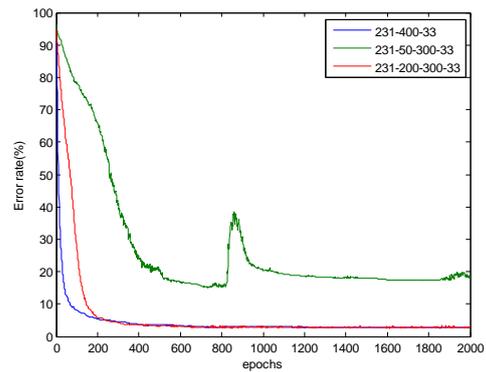


Figure 8: Error rate of different architectures for moment order set to 20

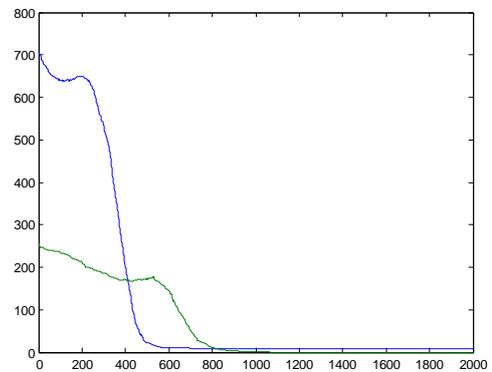


Figure 9: recognition rate of training set of the same architecture and two different samples

The recognition converges faster when the number of samples is great, due to the very small number of training examples.

We believe it is because there is a great level of consistency in how a user draws shape (character). Of course, the more examples, the better is to train the recognizer.

VI. Conclusion

An improved method of construction for handwritten character recognition has been presented. The Legendre moments features used for character recognition are shown to be effective for developing training and test sets which have improved generalization capability without any preprocessing of characters images set.

Further improvements can be made by using more realistic training data and by modifying the hidden layers of the ANN to be sensitive to shifts of characters. The system showed good performance (97%) on a database of 7524 handwritten Amazighe characters.

The results of structure analysis show that if the number of hidden nodes increases the number of epochs (iterations) taken to recognize the handwritten character is also increases. A lot of efforts have been made to get higher accuracy but still there are tremendous scopes of improving recognition accuracy by developing new feature extraction techniques or modifying the existing feature extraction technique

References

- [1] S. Sardar, A. Wahab, Optical character recognition system for Urdu, *International conference on Information and Emerging Technologies*, June 2010.
- [2] M.Soleymani and F.Razzazi, An efficient front-end system for Isolated Persian/Arabic character Recognition of handwritten data-Entry Forms, *International Journal Of Computational Intelligence*, vol 1, pp.193-196,2003.
- [3] B. B. Chaudhuri and U. Pal, A Complete Printed Bangla OCR System, *Pattern Recognition*. Vol.31, 1998, pp.531-549.
- [4] M. K. Hu, Visual Pattern Recognition by Moment Invariant, *IRE Trans. Info. Theory* , vol. IT – 8, pp. 179–187, Feb 1962.
- [5] Bouchra EL BARKANI, *Le choix de la graphie Tifinaghe pour enseigner, apprendre l'amazighe au Maroc: conditions, représentation et pratiques*, Ph.D. thesis, Laboratoire d'Electronique, Signaux-Systèmes et d'Informatique, University of Jean Monnet, Saint-Etienne, France, December 2010.
- [6] M. Ameer, A. Bouhjar, F. Boukhris, A. Boukouss, A. Boumalk, M. Elmedlaoui, E. Iazzi, Graphie et orthographe de l'Amazighe, *Publications de Institut Royal de la Culture Amazighe*, Rabat Maroc, 2006
- [7] M. Outahajala, L. Zenkouar, P. Rosso, A. Martí, Tagging Amazighe with AncoraPipe, *Proc. Workshop on LR & HLT for Semitic Languages, 7th International Conference on Language Resources and Evaluation, LREC-2010*, Malta, May 17-23, 2010, pp. 52-56.
- [8] R. Mukundan, Some Computational Aspects of Discrete Orthonormal Moments, *IEEE Trans. on Image Processing* , Vol. 13, No. 8, pp 1055-1059, Aug 2004.
- [9] J. Haddadnia, K. Faez, and P. Moallem, Neural network based face recognition with moments invariant, in *Proc. IEEE International Conference on Image Processing (ICIP '01)*, vol. 1, pp. 1018–1021, Thessaloniki, Greece, October 2001.
- [10] C.-H. Tech and R.T. Chin, On image analysis by the methods of moments, *IEEE Trans, on Pattern Analysis and Machine Intelligence*, vol. 10, no 4, pp.496-513, 1988.
- [11] S. O. Belkasim, M. Shridhar, and M. Ahmadi, Pattern recognition with moment invariants: a comparative study and new results, *Pattern Recognition*, vol. 24, no. 12, pp. 1117–1138, 1991.
- [12] H. Qjidaa and L. Redouane, Robust line fitting in a noisy image by the method of moments, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1216-1223, 1999.
- [13] H. El Fadili, K. Zenkouar and H. Qjidaa, Lapped block image analysis via the method of Legendre moments, *EURASIP Journal on Applied Signal Processing*, vol. 2003, no.9, pp. 902-913, August 2003.
- [14] X. Zhunang, R. M. Haralick, and Y. Zhao, Maximum entropy image reconstruction, *IEEE Trans. Signal Processing*, vol. 39, no. 6, pp. 1478-1480, 1991.
- [15] M. M. Gupta, L. Jin, and N. Homma, *Static and Dynamic Neural Networks: From Fundamentals to Advance theory*, New Jersey: John Wiley & Sons Inc., 2003.
- [16] A. Khotanzad and J-H. Lu, Classification of Invariant Representations using a Neural Networks, *IEEE Trans. on Acoust., Speech and Signal Process* ., vol. 38, no. 6, pp. 1028-1038, Jun. 1990.
- [17] A.A Zaidan, B.B Zaidan, Hamid.A.Jalab, Hamdan.O.Alanazi and Rami Alnaqeib, Offline Arabic Handwriting Recognition Using Artificial Neural Network, in *journal of computer science and engineering*, vol 1, issue 1, May 2010, pp 55-58.
- [18] Y. Hirose, K. Yamashita, and S. Hijita, Back-propagation algorithm which varies the number of hidden units, *Neural Network*, vol. 4, pp. 61-66, 1991.
- [19] M. hoehfeld, S. E. Fahlman, Learning with limited numerical precision using the cascade-correlation algorithm, *IEEE Tran. On Neural networks*, vol. 3, pp. 602-611, 1992.
- [20] H. El fadili, *Conception d'un système de reconnaissance de forme par combinaison de la méthode des moments avec un classifieur neuronal optimisé par algorithme d'évolution*, Ph.D. thesis, Laboratoire d'Electronique, Signaux-Systèmes et d'Informatique, University of Sidi Mohamed ben Abdellah, Fes, Morocco, 2006.

A Survey of Power Management in Embedded System Using Transistor Sizing

M. Sinduja¹, G. Sathiyabama²

(Department of ECE, Sathyabama University, India)

(Department of ECE, Sathyabama University, India)

ABSTRACT: This paper describes a transistor sizing methodology for both analog and digital CMOS circuits. Various techniques are used for power optimization in CMOS VLSI circuits. Transistor sizing is one of the important techniques for the determination of circuit performance. The aim of the power optimization is to minimize the power and power-delay product or the energy consumption of the circuit. Thus, the main purpose of transistor sizing in CMOS VLSI circuits is to obtain the minimum power dissipation under certain performance requirements. Transistor sizing is based on different algorithms. Analysis is based on some simulation parameter like No. of transistors, power, delay, power delay product, different technologies, aspect ratio. Each of these circuits cell exhibits different power consumption, delay and area in different VLSI technology. This survey paper discusses the different algorithms for transistor sizing analog and digital CMOS circuits and that will be beneficial for the circuit designers.

Key-Words: Algorithm, Delay, Low Power, Optimization, Transistor sizing.

I. INTRODUCTION

Transistor size optimization is a traditional obligation in VLSI (Very Large Scale Integration) design. It is used to improve the performance of a circuit to achieve a design goal in a specific technology. This design goal can either be boosting operating speed, lowering power consumptions or lowering area requirements. Design of VLSI circuits involves tradeoff between high speed (low delay) and low power. Design requires optimum balance between delay and power, in smaller constituent circuits and also in the larger circuit. This can be achieved by careful iterative optimization of transistor sizes. In this paper, the net list of the circuit is determined by adjusting the length and width of the MOS transistors. The figure of merit depends in a complex way on the individual size of the transistors. Changing transistor sizes in a circuit often leads to surprising results, which are not easily predicted in the low power design embedded system design.

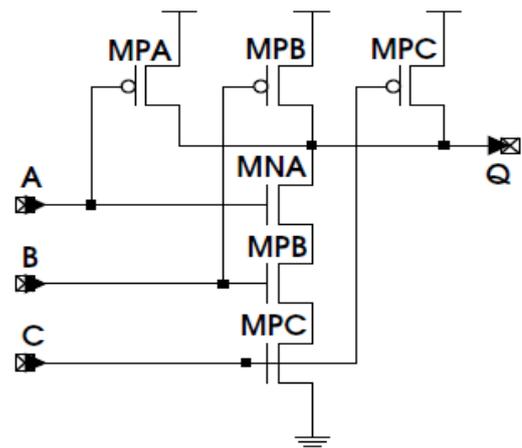
The different algorithms have been presented for transistor sizing. The remainder of the paper is organized as follows. In section [2.1] the Monte Carlo and Genetic algorithm optimize the three digital CMOS circuits and in section [2.2] annealing and ACM compact model are discussed. Then after a simple declaration of Genetic Algorithm on consuming energy for basic arithmetic circuits, the Genetic circuit optimizer is proposed in section [2.3]. In section [2.4] optimisation techniques using geometric programming are discussed. Comparison of different transistor sizing analog and digital CMOS circuits

performance and results are shown in section [3] and the conclusion is presented in section [4].

II. LITERATURE SURVEY

2.1 MONTE CARLO AND GENETIC ALGORITHM

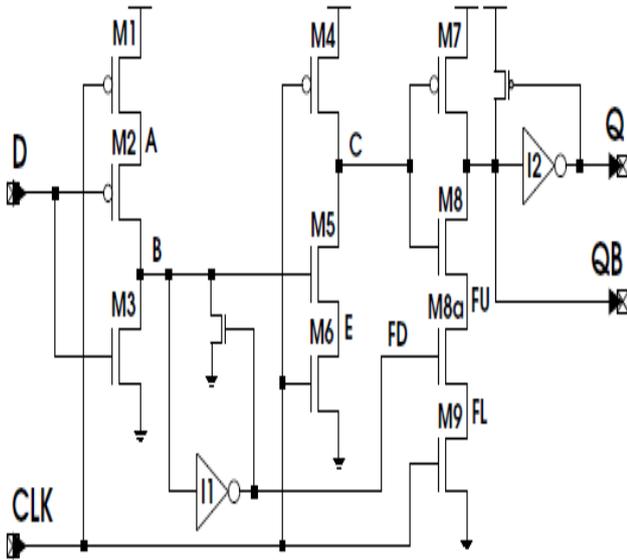
This paper [1] presents the Stochastic methods for transistor size optimization of CMOS analog and digital circuits. The performance of CMOS circuits depends on transistor sizes. The transistor sizes of digital circuits are optimized using Monte Carlo and Genetic Algorithm (GA), and also comparison is done between the two algorithms. GA is found to be more effective for larger circuits while for smaller circuits Monte Carlo optimization performs better than GA. Even nowadays, GA is one of the most used algorithms for optimization problem. This paper reviews three techniques a standard optimizer, a Monte Carlo scheme and a method based on genetic algorithms combined with very accurate spice simulations to automatically optimize transistor sizes of three digital CMOS circuits as three input NAND gate, Dynamic edge triggered DFF, Complementary pass gate full adder (CPLFA). This paper improves the performance to achieve a design goal. This design goal can either be boosting operating speed, lowering power consumption or lowering area requirements.



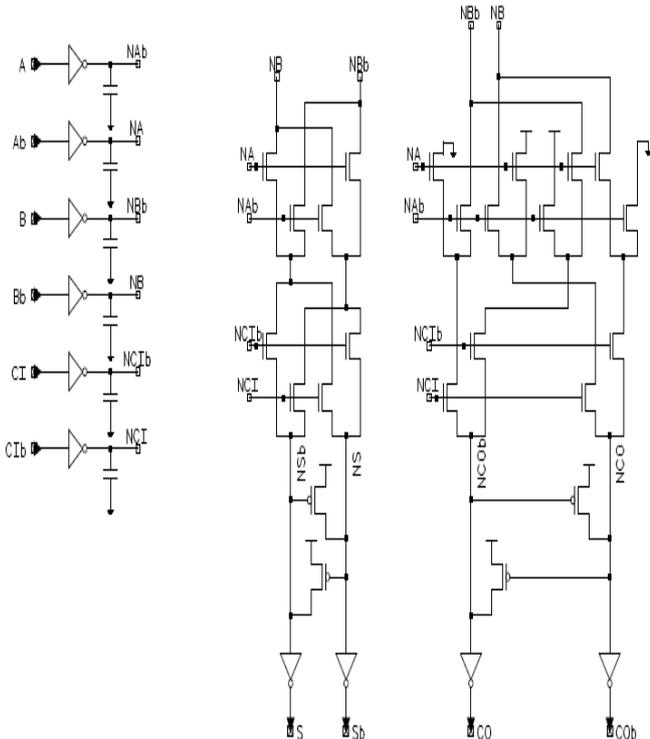
“Fig 1” 3-Input CMOS NAND Gate

The Monte Carlo algorithm used in this paper results in minimum duration because the step size is first set to four times the minimum feature size and then to twice the minimum feature size and the minimum for the last third. The genetic algorithm is based on ranking selection, uniform cross over and the optimal mutation rate. In this paper the optimization of transistor sizes can be accurately performed by stochastic optimization. The paper compared

four stochastic optimization methods on three CMOS subcircuits of different complexity (6, 16, 64 MOSFETS).



“Fig 2” Dynamic Edge triggered DFF

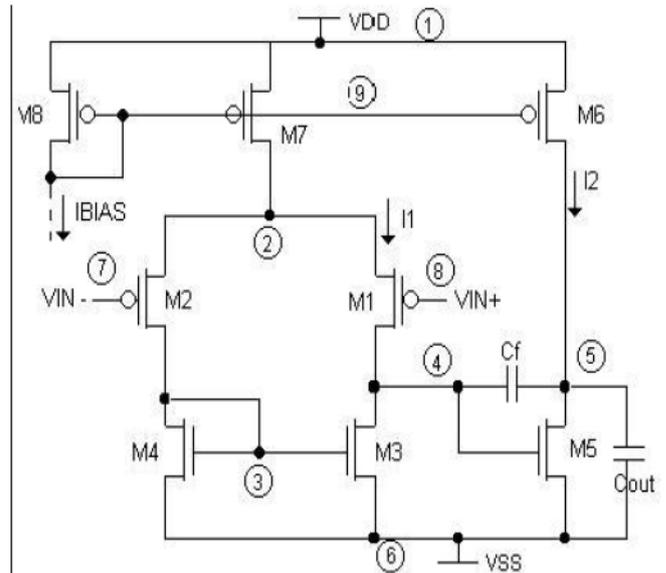


“Fig 3” Complementary Pass Gate Full Adder (CPLFA)

On the smaller two subcircuits, the Monte Carlo method has been found to yield better results than the genetic algorithms investigated. On the largest of the three circuits, the genetic algorithms were found to yield significantly smaller variances provided they were allowed to run for a significant number of generations which suggests that the monte carlo methods tends to end up in a suboptimal solution more frequently as problem size increases.

2.2. ANNEALING AND ACM COMPACT MODEL

The paper [2] described a transistor sizing methodology for analog CMOS circuits that combines the physics-based *gm/ID* characteristics provided by the ACM compact model and the simulated annealing technique for the circuit optimization. This methodology exploits different transistor widths and lengths and provides good solutions in a reasonable CPU time, with a single technology dependent curve and accurate expressions for transconductance and current in all operation regions. The advantage of constraining the optimization within a power budget is of great importance for the designer. The optimization results obtained for the design of a two-stage operational amplifier. This paper discussed a methodology for analog design automation that combines the simulated annealing optimization technique, a physics-based transconductance to-current ratio characteristics and the electrical simulation. The



“Fig 4” Schematics of a two-stage operational amplifier

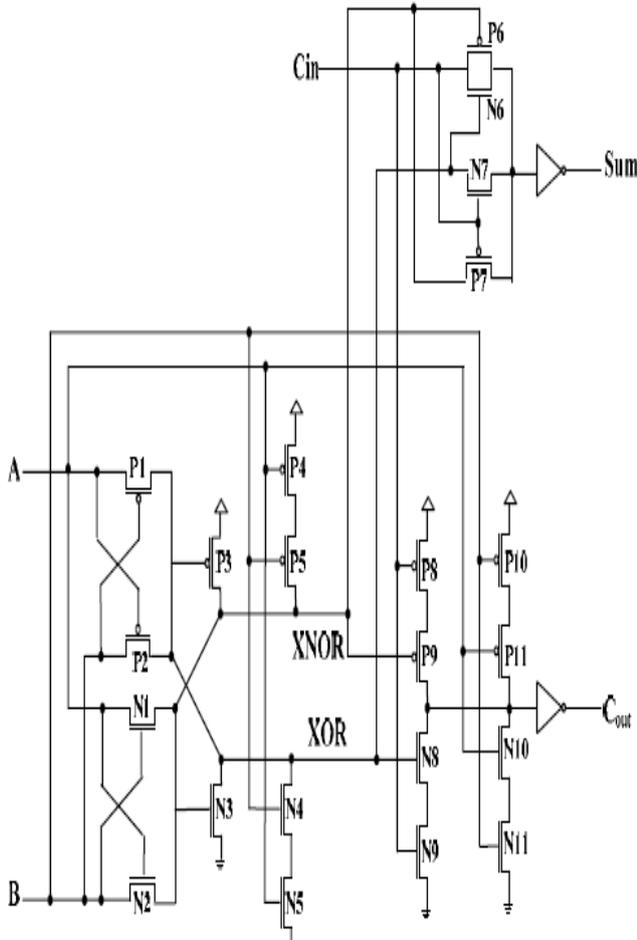
Synthesis of a two-stage operational amplifier is shown in order to demonstrate the capabilities of the methodology. The paper describes simulated annealing process in which metal cools and freezes into a minimum energy crystalline structure. This algorithm employs a random search which does not only accept solutions that decrease the objective cost function *f_c* but also some changes that increase it. This paper exploits different transistor lengths and provides an acceptable solution in a reasonable CPU time. It is a simple sizing method based on transistor inversion coefficient which is calculated by curve *gm/ID* versus *I*.

2.3. BASED ON GENETIC ALGORITHM

In this paper, transistor sizing is very important for determination of the circuit performance. This paper [3] is based on the methodology of genetic algorithm for energy consumption optimization. Simulation results in this paper show that, compared with the hybrid tree structure, the GA transistor sizing exhibits better simplicity, initial values, independency, optimization parameter independence and short runtime. The average improvement in PDD is 7% for

XOR/XNOR circuits and 17% for full adder. In this technique, all the circuits are simulated using Hspice circuit simulated based on two techniques 0.13um and 0.18um based on BS1M3V and TSMC model.

The main objective of annealing and compact model is to minimize the power delay product on the energy consumption. The energy consumption was optimized. The transistor size run time is reduced by using Genetic algorithm in transistor sizing optimization. Finally compared



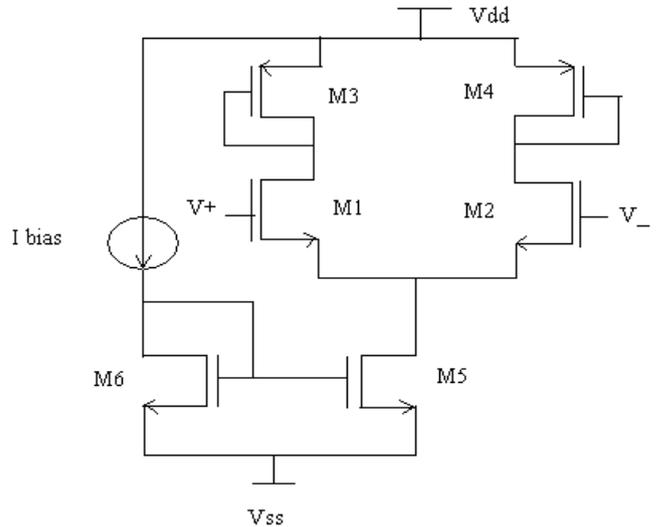
“Fig 5” The New HPSC Full adder

to the hybrid tree structure algorithm, the PDP characteristics are improved by 7% in XOR/XNOR circuit and 17% in full adder for the supply voltage range of 0.6 to 0.2 V. The saving in PDP increases in the larger circuit.

2.4. OPTIMIZATION USING GEOMETRIC PROGRAMMING

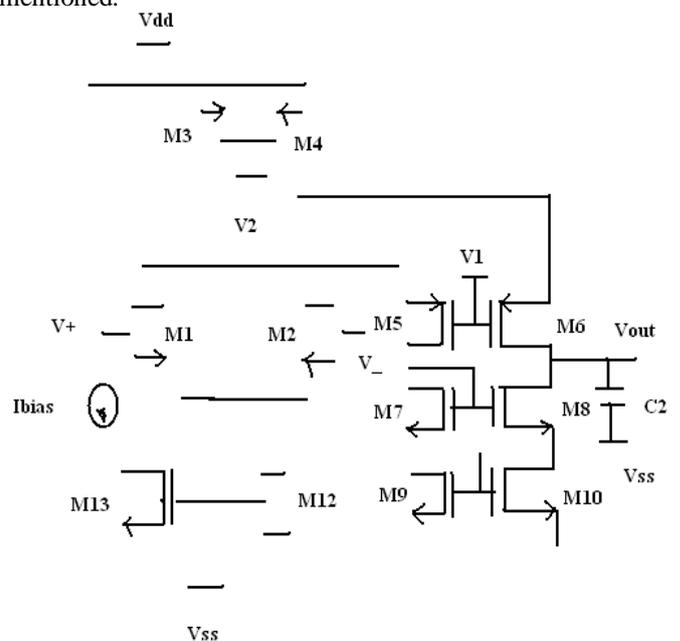
This paper describes a design of folded cascode operational amplifier with power optimization using geometric programming. The basic idea of the folded-cascode operational amplifier is to apply cascode operational amplifier transistors to the input differential pair. In this study, the main concepts of power optimization using geometric programming are introduced. Geometric Programming, a method for optimizing and automating component and transistor sizing will apply to the power optimization of Folded Cascode operational amplifier. The simulation of the folded cascode circuit is done using TSPICE simulation tool and 0.035 μm CMOS technology is

used. Design of the circuit and simulation results are presented demonstrating the amplifier power optimization. In this paper we introduced a new method, geometric programming, for determining the component values and transistor dimensions for CMOS op-amps. The method handles a wide variety of specifications and constraints, is extremely fast, and results in globally optimal designs.



“Fig 6” Differential amplifier with cascode arrangement

This paper describes geometric programming, for circuit sizing and optimization. Initially our work describes the design of differential amplifier with cascode arrangement and then the design of folded-cascode operational amplifier. The simulation folded cascode op-amp circuit is done using TSPICE simulation tool and the T-spice parameters are mentioned.



“Fig 7” Folded Cascode operational amplifier

Simulated result shows that this op amp exhibits the average power is 1.007X10-13 that is 100 fW or 0.1 nW operating at 100 KHz frequency. The Geometric Programming is successfully applied and the optimized

average power is obtained 9.0658×10^{-14} that is 90 fW or 0.09 nW operating at 100 KHz frequency.

“Table 1” Comparison of different transistor sizing analog and digital CMOS circuits

S.No	Transistor Sizing	Algorithm	Circuit	Technology	Power	Run time	No of Transistors
1.	Robert Rogenmoser, Hubert Kaeslin, Tobias Blicke in 1996	Schoastic Monte Carlo and genetic algorithm	NAND3	1.0 μ m	3.0	Less	6
			DFF	1.0 μ m	5.8		12
			CPLFA	0.6 μ m	36.7		26
2.	Alessandro Girandi and Sergio Bampi in 2006	Annealing and ACM compact model	OPAMP	0.35 μ m	-	More	8
3.	M. Grailoo, T. Nikoubin, K. Navi in 2009	Genetic algorithm	HPSC Full adder (XOR/XNOR gate)	0.13 μ m	0.406	More	18
				0.18 μ m	1.292		
4.	Amol W. Pardhi and Dr.A.Y. Deshmukh in 2012	Geometric programming	Folded cascode operational amplifier	0.035 μ m CMOS	9.0658×10^{-14}	Less	12

III. CONCLUSION

Transistor sizing is an effective technology for reducing the power, delay in the analog and digital CMOS circuits of an embedded system. In this way different technique such as genetic algorithm, geometric programming, Monte Carlo algorithm are reviewed and compared the performance from the most recent published research work. Different algorithms are discussed to reduce the size of the transistors in which genetic algorithm is more efficient and it leads to better circuit performance in the case of power consumptions. The size of the transistor size is reduced under 180° and further proceeding should be carried our below 100°. By using the analysis of different algorithms the low power circuit can be implemented in embedded system.

REFERENCES

[1] S.S. Sapatnekar, V.B. Rao, P.M. Vaidya and Sung-Mo Kang, "An exact solution to the transistor sizing problem for CMOS circuits using convex optimization," *Computer-*

Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol.12, no.11, pp.1621-1634, Nov. 1993.
 [2] K. Deb, "An introduction to genetic algorithms," *Sadhana*, Vol.24, Parts 4 & 5, August & October 1999, pp. 293-315.
 [3] Haupt, R.L., Haupt, S.E, 2004, "Practical Genetic Algorithms", *John Willey & Sonc, Inc.*
 [4] Varun Aggarwal and Una-May O'Reilly, "Design of Posynomial Models for Mosfets: Symbolic Regression Using Genetic Algorithms", *Genetic Programming: Theory and Practice IV* : pp. 219-236, 2006.
 [5] F. Aezinia, A. Afzali-Kusha and C. Lucas, "Optimizing High Speed Flip-Flop Using Genetic Algorithm," *Circuits and Systems, APCCAS 2006, IEEE Asia Pacific Conference on* , pp.1787-1790, 4-7 Dec. 2006.
 [6] Y. Chien-Cheng, "Low-Power Double Edge-Triggered Flip-Flop Circuit Design," *Innovative Computing Information and Control, 2008, 3rd International Conference on*, pp.566, 18-20 June 2008.
 [7] R.A. Thakker, M.S. Baghini and M.B. Patil, "Low-Power Low- Voltage Analog Circuit Design Using Hierarchical Particle Swarm Optimization," *VLSI Design 2009, 22nd International Conference on*, pp.427-432, 5-9 Jan. 2009.

- [8] M.J.P. Jalali and A. Rezaee, "Power optimization in 70nm technology," *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, vol.3, pp.V3-136-V3-139, 16-18 April 2010.
- [9] T. Nikoubin, S. Pouri, P. Bahrebar and N. Navi, "A New Transistor Sizing Algorithm for Balanced XOR/XNOR Circuits," *CSICC. Faculty of Electronical and Computer Engineering, SBU, Tehran, Iran February 20-22, 2007*, [12th International CSI Computer Conference, CSICC 2007, Tehran, Iran] (in Persian).
- [10] T. Nikoubin, P. Bahrebar and N. Navi, "A New Design for XOR/XNOR Circuits, Based on High-Speed and Low-Power Arithmetic Circuits used Signal Processing", *BEC Magn. Iran, vol. 1, pp. 203-209*, [second Broadcast Engineering Conference. IRIB University, Iran, 17-19 November 2007] (in Persian).
- [11] JH Yu, ZG Mao, "A design method in CMOS operational amplifier optimization based on adaptive genetic algorithm", *WSEAS Transactions on circuit and systems*, vol. 8, pp 548-558, 2009.
- [12] Er. Rajni, "Design of High Gain Folded-Cascode Operational Amplifier Using 1.25 um CMOS Technology" *International Journal of Scientific & Engineering Research* Volume 2, Issue 11, November-2011

Magic Square

Dr. Mangala Gurjar

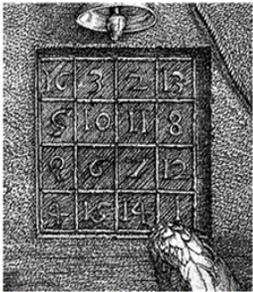
St. Xavier's College, Mumbai 400001

Abstract: In this paper we will discuss an algorithm to generate a magic square of any odd order. We will also see computer programs for this algorithm in the programming languages C which is a popular programming language and in BASIC which is a higher level language suitable to translate algorithms. The programs can be run on computers to generate a magic square of order $n \times n$ for any odd value of n .

Key words: Algorithm, Array, BASIC Program, C-program, Magic square.

I. Introduction

Magic squares have fascinated humans for many centuries. Following is a magic square created in 1514 by Albrecht Durer.



It looks like this.

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

And what is the magic?
It is the number 34!

This number is the sum of the various fields within the magic square.

Since then many magic squares were made. The latest one is a 3125 x 3125 magic square made by Mr. Suresh Sutar of Kolhapur, Maharashtra in October 2011.

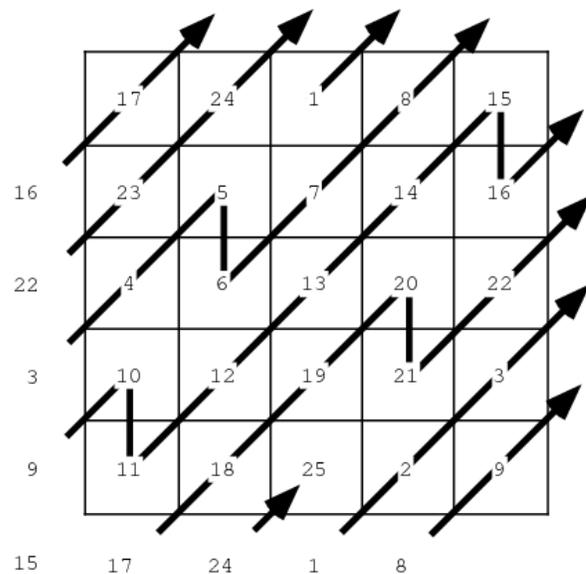
II. Algorithm

Following is an algorithm that can generate a magic square of size $n \times n$ for any odd value of n .

- Step 1 Input an odd number n .
- Step 2 Create an $(n+1) \times (n+1)$ integer array having all entries equal to zero.
- Step 3 Set m equal to 1. Row number i to 1 and Column number j to $(n+1)/2$
- Step 4 While m is not bigger than n^2 repeat the following steps up to step 11.
- Step 5 Insert m in the (i,j) th place.

- Step 6 Decrease i by 1 and increase j by 1. Increase m by 1.
- Step 7 If $i = 0$ and j is not equal to $n+1$ then change i to n .
- Step 8 If $j = n+1$ and i is not equal to zero then change j to 1.
- Step 9 If $i = 0$ and $j = n+1$ then change i to 2 and j to n .
- Step 10 If the (i,j) th entry is non zero then increase i by 2 and decrease j by 1.
- Step 11 Increase m to $m+1$.
- Step 12 Print the two dimensional array.

The schematic diagram for the above algorithm for $n = 5$ is given below.



III. C-Program

Following is the c implementation of this algorithm

```
# include "stdio.h"
# include "conio.h"
int main (int n)
{
printf("Enter\n");
scanf("%d" , &n);

int index = n;
int a[index+1][index+1];
int i,j,m;

for(i=1 ; i<=n ; i++)
    for(j=1 ; j<=n ; j++)
        a[i][j]=0;

m = 1; i = 1; j = (n+1)/2 ;
While ( m <= n*n)
{
```

```

a[i][j] = m ; m++ ; i = i - 1 ; j++ ;
If (i ==0 && j<>n+1)then i = n ;
If (i <> 0 && j == n+1) then j = 1;
If(i == 0 && j == n+1) then {i = 2; j = n};
If a[i][j]<>0 then {i = i+2 ; j = j-1}
}
for(i=1; i<=n ; i++)
    for(j=1 ; j<=n ; j++)
    {
    printf("%d", a[i][j]);
    if(j=n) printf("\n");
    }
getch();
return 0;
}

```

IV. BASIC-Program

Following is the Basic program for the same algorithm.

```

10 REM Magic Square
20 INPUT "The dimension"; N
30 DIM A(N+1, N+1)
40 FOR I = 1 To N+1
50 FOR J = 1 TO N+1
60 LET A(I, J) = 0
70 NEXT J
80 NEXT I
90 LET I = 1 : J = (N + 1)/2
100 FOR M = 1 TO N*N
110 LET A(I, J) = M
120 LET I = I - 1 : J = J + 1
130 IF I = 0 AND J <> N + 1 THEN I = N
140 IF J = N + 1 AND I <> 0 THEN J = 1
150 IF I = 0 AND J = N + 1 THEN I = 2 : J = N
160 IF A(I, J)<> 0 THEN I = I + 2 : J = J - 1
170 NEXT M
180 FOR I = 1 TO N
190 FOR J = 1 TO N
200 PRINT A(I, J) ;
210 NEXT J
220 PRINT
230 NEXT I
240 END

```

V. Conclusion

An $n \times n$ magic square for any odd large value of n can be formed using the above simple computer program. This eliminates claims made by lay people of having set a new record of making the largest magic square. Similar program can be written for even values of n also.

REFERENCES

- [1] R.G. Dromey, How to Solve it by computers
Prentice-Hall India.
- [2] Kenneth H. Rosen : Discrete Mathematics and Its
Applications, McGraw Hill

The influence of physico-chemical parameters of fountain solution on print quality

Ivana Oros

University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design,
Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia

ABSTRACT: The influence of pH, conductivity and Ca^{2+} ion concentrations of fountain solutions on the quality of tone, dot, line and text was investigated. Physico-chemical parameters of fountain solutions were analyzed by standard methods. Image quality assessment (IQA) was carried out using the ISO methodology and ImageJ software. The obtained data of the mechanical and optical tone value, dot circularity, line and text raggedness were compared and correlated with the physico-chemical parameters of fountain solutions. The results of tone and dot quality indicated the great influence of pH, conductivity and Ca^{2+} ion concentrations of fountain solutions on the formation and circularity of magenta dots. The highest mechanical and optical tone value increases were observed in the mid tones of magenta color. The mutual nonlinear dependence, described by second-order polynomial, was existed between magenta dot circularity and the printed sheets. The results of line and text raggedness indicated that magenta lines and text and analyzed physico-chemical parameters of fountain solution were not in the mutual dependence.

Keywords: Ca^{2+} ion, conductivity, fountain solution, pH, print quality, sheet-fed offset

I. INTRODUCTION

Offset lithography relies on the numerous key parameters, and among them the fountain solution plays an important role. Indeed, its main tasks are to:

- wet and desensitize the non-image areas of the printing form;
- maintain hydrophilic character of the non-image areas;
- prevent the acceptance of ink on the non-image areas;
- maintain the working properties of ink;
- aid in proper blanket release, which reduce piling on the blanket [1-3].

The fountain solution, as one of the major component of the offset lithographic process, contains water, wetting agent (isopropyl alcohol), buffer, desensitizing agent (gum arabic), corrosion inhibitor, biocide and additives. Fountain solution effectiveness depends on properly mixing and maintenance of all fountain solution components. Accurately-measured pH value and conductivity of fountain solution are both essential for quality printing. While these two parameters are basically independent, each one provides vital information about the used water and the prepared fountain solution in the fountain solution dosing system. Fountain solutions are normally buffered at specific pH value, which means that pH of fountain solution decreases as the buffer concentration increases. When the certain pH value of fountain solution is reached the additional amount of buffer does not affect its changes. pH value increases when the buffer capacity

weakens and then fountain solution becomes more alkaline. Conductivity, as other important parameters of the fountain solution, accurately determines the initial fountain solution strength and indicates its contamination during the printing. The ink and paper, as a primarily alkaline-based matter, contaminate the fountain solution by changing its pH value to alkaline and weakening the buffer capacity. This fountain solution affects the final print quality and leads to scumming, toning, tinting, etc. [2, 4-6].

Also, many printing problems could be directly traced to the increased presence of calcium in the fountain solution. The common sources of calcium are:

(i) **water** - most calcium compounds present in water are insoluble in acid-based fountain solution. Calcium carbonate slowly reacts with the acid-based fountain solution by releasing calcium ions, which can combine with citrate and phosphate ions from buffer, and precipitate out on rollers in the fountain system as a white haze or as hard, rock-like deposits. Also, carbonates increasing pH value of the fountain solution and reducing buffer capacity. Conductivity usually, but not always, increases with increasing of water hardness and alkalinity;

(ii) **ink** - calcium-based red pigments are mostly used in the formulation of magenta inks. The constant mixing of fountain solution and magenta ink during the printing causes calcium to be absorbed into the fountain solution. Once there, it is free to react with other present ions, and can form deposit, insoluble compounds. Fountain solutions in press units with magenta ink may also experience sharp upward rises in pH and conductivity within a short period of time [7, 8]; and

(iii) **paper** - in recent years most paper manufacturing has switched from an acid to an alkaline paper production process [7, 8]. Alkaline paper is made under slightly alkaline (pH 7.5 to 8.5) conditions, uses an organic sizing, and accommodates higher amounts of filler - usually 15 to 25% calcium carbonate [9, 10]. Improper or insufficient sizing or poor coating can lead to piling and the release of calcium carbonate onto the blankets. Then, it can build-in into the ink and dampening roller trains via the printing forms in the fountain solution [7, 8].

Therefore, the prevention of the negative effects of pH, conductivity and calcium concentration on the final print quality, as essential, implies the monitoring of these parameters during the printing and requires the application of the image quality assessment.

A common way to analyze the print quality is to objective assess the tone and color with light-reflection measuring device. Although the tone and color, as the quality parameters, are easily perceptible, they are not sufficient to determine the print quality. Therefore, the analysis of print quality must include and parameters of line

and dot quality. The International Standards Organization was developed the ISO 13660 standard [11], which in accordance with ISO 19751 [12] provides definitions of 14 different printed image's attributes that help analyze the print quality. The attributes are categorized in two groups [13]:

1. Area attributes: darkness, background haze, graininess, mottle, extraneous marks, and background voids.
2. Character and line attributes: blurriness, raggedness, line width, darkness, character contrast, fill, extraneous marks, character field, and background haze.

The objective of this investigation was to monitor the parameters of the fountain solution quality (pH, conductivity and Ca^{2+} ion concentration) during the printing of 24,000 sheets and to determine their possible correlation with the print quality parameters (mechanical and optical tone value increase, dot circularity, line and text raggedness).

II. MATERIALS AND METHODS

2.1 Offset printing process

A four-color offset printing press Heidelberg SM HD102VP, which is installed in printing facility Forum, Novi Sad, and printed color in the sequences CMYK, was used in the printing process. Also, printing forms (KODAK, Germany) with CMYK color strips including the patches of 0, 2, 4, 6, 8, 10, 20, 30, 40, 50, 60, 70, 80, 90, 92, 94, 96, 98 and 100% tone value, horizontal and vertical lines (1, 1.5 and 2pt) and Times and Arial text (4, 5, 6, 7, 8, 9, 10, 11 and 12pt font size) was used, Fig. 1.

For this investigation thirteen printed sheets (samples 1-13) were collected from the delivery unit during the printing of 24,000 sheets. Samples 1 and 13 were taken at the beginning and at the end of a print run, respectively. Samples 2-12 were every 2,000 printed sheet during a print run. Sample 7 was taken at the half of a print run, i.e. when 12,000 sheet was printed and it is used as a reference.

2.2 Fountain solution

Fresh fountain solution was prepared in printing facility Forum by mixing of 2% buffer (P56 Alkopufer, Cinkarna Celje, Slovenia), 12% isopropyl alcohol (P43 ISO fount, Cinkarna Celja, Slovenia) and 86% tap (untreated) water. Thirteen samples of fountain solution (samples 1-13) were taken for the monitoring of pH, conductivity, total dissolved solids and Ca^{2+} ion concentrations during the printing of 24,000 sheets. Each sample was followed by the sample of printed sheets. Sample 7 of fountain solution is used as a reference.

2.2.1 Analysis of tap water and fountain solution

The preparation of fountain solution was included the monitoring of pH, conductivity and total dissolved solids of tap water in printing facility Forum using HI 98129 instrument (USA) with accuracy of measurement: ± 0.05 pH, conductivity $\pm 0.02 \mu\text{S}/\text{cm}$ and total dissolved solids ± 0.02 ppm. The same instrument was used for measurement of the mentioned parameters in fountain solution samples. The Ca^{2+} ion concentration in samples of tap water and fountain solution was determined by Inductively Coupled Plasma with Mass Spectrometry (ICP-MS), using a PerkinElmer Elan 5000 mass spectrometer.

Water hardness is determined by complexometric titration of water samples with a solution of complexone III and Eriochrome Black T indicator.

Surface tension of fountain solution is measured by using stalagmometer. The density of fountain solution samples was firstly measured by pycnometer, and then the number of drops of each fountain solution sample, which flows from the constant volume of stalagmometer, was determined. Surface tension of fountain solution is calculated by the following formula (1):

$$\gamma_i = \gamma_0 \frac{n_0 \rho_i}{n_i \rho_0} \quad (1)$$

where is: γ_i - surface tension of fountain solution, γ_0 - surface tension of water, n_i - number of fountain solution drops, n_0 - number of water drops, ρ_i - density of fountain solution and ρ_0 - density of water [14].

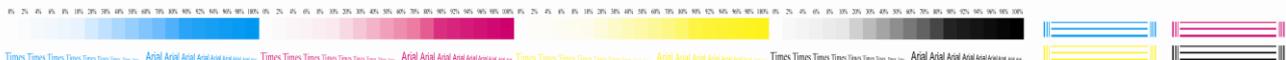


Fig. 1 Color strip with CMYK tone values, lines and text

2.3 Paper and ink

A glossy coated paper (BIOGLOSS, B&B Papirnica, Vevče) with the characteristics presented in Table 1, defined as Type I in ISO 12647-2: 2004 [15] were printed with sheet-

fed offset ink set (Inkredible RAPIDA F 10 RP, Huber group, Germany) with general chemical composition presented in Table 2.

Table 1 Characteristics of Biogloss paper [16]

Property	Value	Unit	Method	Tolerance
Grammage	90	g/m^2	ISO 536	+/- 4%
Caliper	65	μm	ISO 534, single sheet measurement	+/- 5%
Brightness	95	%	ISO 2470, R457 D65	+/- 2%
Opacity dry	91	%	ISO 2471	+/- 2%
Gloss Lehmann	72	%	TappiT 480, 75°	+/- 5%
Picking resistance dry (IGT)	60	cm/sec	ISO 3783 IGT, L3803	min.
Relative humidity	45	%	Hygromer Rotronic	+/- 5% at 23 °C
pH value	7	-	ISO 6588	+ 1%

Table 2 General chemical composition of Inkredible RAPIDA sheet-fed offset ink [17]

Component	mass %
Pigment (organic)	10 - 35
Carbon black	0 - 20
Resin	20 - 35
Vegetable oil	15 - 20
Mineral oil	15 - 20
Additives	>10

Literature data [18-20] indicated that the commonly used pigments in the production of CMYK sheet-fed offset inks are copper phthalocyanine or cyan (CI Pigment Blue 15:3), lithol rubine or magenta (CI Pigment Red 57:1), diarylide yellow or yellow (CI Pigment Yellow 13) and carbon black or black pigment (CI Pigment Black 7), respectively.

The chemical structures of these pigments, Fig. 2, show that only lithol rubine or magenta pigment in its structure contains calcium. Therefore, only the magenta ink samples (two samples) were taken in order to determine the presence of calcium.

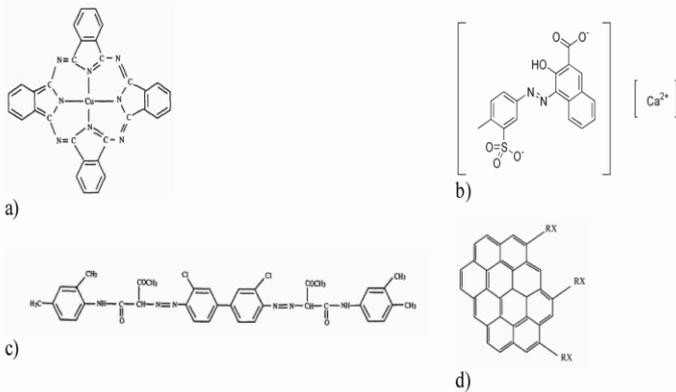


Fig. 2 Chemical structure of pigments for CMYK process inks: a) copper phthalocyanine - cyan, b) lithol rubine - magenta, c) diarylide yellow - yellow and d) carbon black - black [19, 20]

Sample 1 was the reference magenta ink sample, taken from the original ink packaging prior the printing process, while magenta ink sample 2 was taken from the ink unit at the half of a print run, i.e. when 12,000 sheet was printed. Sample 2 of magenta ink was taken with the sample 7 of fountain solution.

Also, the sample of coated paper was taken for the analysis of calcium presence.

2.4 Image quality assessment (IQA)

The IQA was carried out using the ISO 12647, ISO 13660 and ISO 19751 methodology in order to analyze the quality of tone, dot, horizontal and vertical line, serif and sans-serif text style on thirteen printed sheets (samples) during the changes in fountain solutions quality. For this purpose the device: Spectrophotometer SpectroPlate (TECHKON GmbH, Germany), Spectrophotometer SpectroDens (TECHKON GmbH, Germany), flat-bed scanner CanoScan 5600F (Canon Inc., Canada), and software Adobe Photoshop CS3 and ImageJ (version IJ 1.45m) were used.

2.4.1 Analysis of tone value

As the final print quality is the most important, it was necessary to evaluate how the changes of the fountain solution quality (pH, conductivity and Ca²⁺ ion concentration) influence on the dot formation and tone value increase. Therefore, the mechanical tone value increase was firstly measured after the printing in order to establish a reference point for the evaluation of the optical tone value increase on the printed sheets. Both, the mechanical and optical tone values were generated from the spectrophotometric measurements using the spectrophotometers SpectroPlate and SpectroDens on the above mentioned patches.

2.4.2 Analysis of dot, line and text

For achieving the quality of the continuous tone the printing of the smallest elements (screen dot, line and text) is essential. Attributes such as line and text raggedness and dot circularity were of interest for this investigation, since they have an obvious and major influence on the quality of any print. The analysis required the usage of 'reference prints' which were visually and by software compared to the actual samples in order to get the significant quantitative data. As the actual samples it was used:

- (i) the microscopic images of 30% CMYK tone value recorded by the spectrophotometer SpectroPlate for the analysis of dot circularity; and
- (ii) the samples scanned by using scanner CanoScan and software Adobe Photoshop CS3 (1200 dpi scanning resolution) for the analysis of line and text raggedness.

After this procedure, the print defects are easily detected and quantified by the ImageJ software.

III. RESULTS AND DISCUSSION

3.1 Analysis of tap water and fountain solution

In order to obtain the high-quality results during the offset printing process it is essential to accurate control the physico-chemical parameters of tap water and fountain solution. The values of conductivity and total dissolved solids (TDS) (Table 3) indicated that used tap water contains the significant amounts of dissolved ions, which is confirmed by hardness of 11 °dH. Although the used water is hard, it is suitable for the offset printing process.

Table 3 Physico-chemical parameters of tap water

Parameter	Value	Unit
pH	7.49	-
Conductivity	497	μS/cm
Hardness	11	°dH
TDS	196.9	mg/l CaCO ₃
Ca ²⁺ ion concentration	249	ppm
	155.2	mg/l

The investigation showed that the values of pH, conductivity, total dissolved solids and Ca²⁺ ion concentrations of fountain solution samples significantly varied during the printing (Figs. 3-7) with the increasing tendency of all parameters in sample 7 (when 12,000 sheet was printed). The explanation for those phenomena is the fact that conductivity and total dissolved solids rises as the press runs due to increased contamination of fountain solution by inks, paper dust, metals particles from printing form and atmospheric gases. Whereas, increasing of pH value of fountain solution is caused by buffer capacity weakening during the interaction of fountain solution with paper, ink, plate coatings or printing form. Exactly higher print speed and interaction between the used materials during the reproduction process gradually increase the fountain solution contamination in Baldwin system by causing weakening of buffer capacity which was reflected through the slightly increasing of pH, conductivity, total dissolved solids and Ca²⁺ ion concentrations in samples 1 to 6. Significantly increasing of all mentioned parameters are observed at the half of a print run (12,000 sheets), sample 7, when buffer totally disappear from fountain solution. After that, the decreasing trend of pH, conductivity, total dissolved solids and Ca²⁺ ion concentrations is observed in samples 8 to 13 due to the new quantities of buffer and tap water was dosed in Baldwin system, which significantly diluted the contaminated fountain solution and allowed its further usage in the printing process.

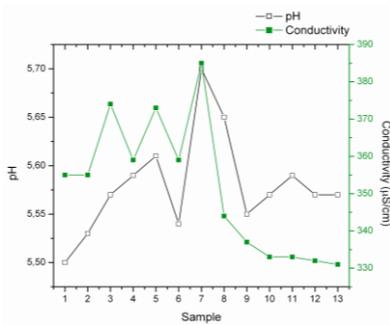


Fig. 3 A correlation between pH and conductivity

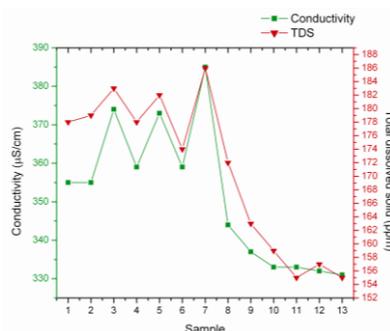


Fig. 4 A correlation between conductivity and TDS

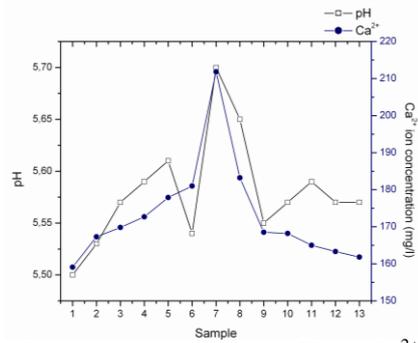


Fig. 5 A correlation between pH and Ca²⁺ ion

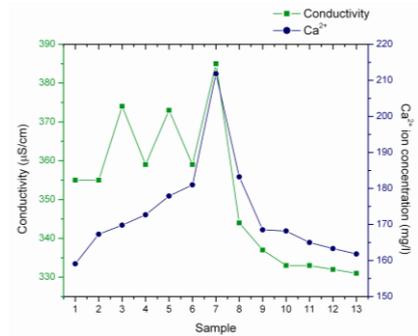


Fig. 6 A correlation between conductivity and Ca²⁺ ion

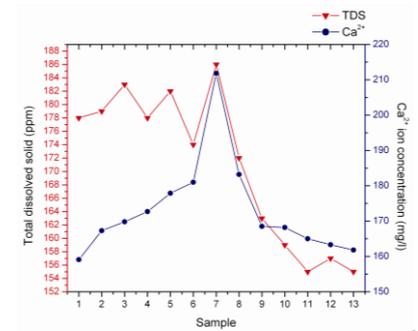


Fig. 7 A correlation between TDS and Ca²⁺ ion

The surface tension (Fig. 8) has the similar trend, and its value was relative constant in range from 38.4 to 41.6 mN/m, before and after sample 8. Sample 8 indicates that an alcohol control (Balcontrol) system, which is installed on Baldwin dosing system, dosed the new quantities of isopropyl alcohol which currently decreases the surface tension in fountain solution.

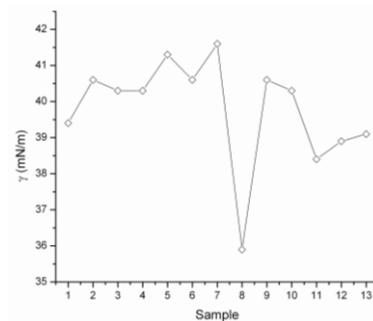


Fig. 8 A surface tension of fountain solution

As calcium carbonate can affect the quality of fountain solution, it was necessary to examine the reasons of pH, conductivity, total dissolved solids and Ca²⁺ ion

concentrations changing in the fountain solution samples by the analysis of used paper and magenta ink.

3.2 Analysis of paper and magenta ink

The investigation primarily involved the characterization of used paper in order to examine the filler type and the possible presence of calcium carbonate.

SEM analysis of used Biogloss paper (Fig. 9) by including the paper surface shown that inorganic filler stands out as bright sparsely distributed particles against the pulverized background of paper, i.e. it was not observed the network-like structure of paper with fiber-fiber bond and crossing.

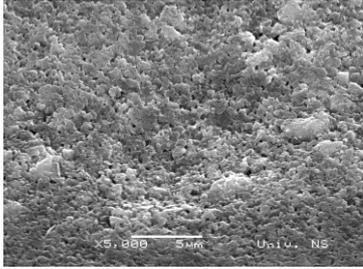


Fig. 9 SEM image of used Biogloss paper (magnification 5,000 x)

As can be seen from Fig. 9 the particles of kaolin (which tend to be platy) and coarsely ground natural calcium carbonate (GCC) (which has irregular edges) are present in Biogloss paper.

In order to accurately determine the elemental chemical composition of the paper sample, i.e. which fillers are presented in Biogloss paper, the analysis by the energy dispersive spectroscopy (EDS) was carried out. Obtained EDS spectra (Fig. 10) shows that calcium carbonate is dominant in Biogloss paper. Magnesium with aluminum, silicon and iron suggests that kaolin is present. While calcium alone (or with carbon and oxygen) suggests that calcium carbonate is present. Thus, the EDS analysis confirmed the presence of kaolin and calcium carbonate as the fillers in Biogloss paper. Proportions of all inorganic elements in Biogloss paper sample are given in Table 4.

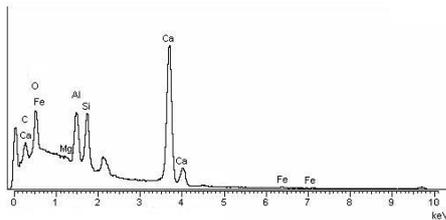


Fig. 10 EDS elemental spectra of Biogloss paper

Table 4 Chemical composition of Biogloss paper

Element	mass %
C	21.25
O	47.27
Mg	0.33
Al	4.99
Si	5.27
Ca	20.30
Fe	0.38

Also, magenta ink samples were examined using EDS in order to control the concentration of calcium ions in

magenta ink during the printing as well as to confirm the increased Ca²⁺ ion concentrations in sample 7 of fountain solution (when 12,000 sheet was printed).

The obtained EDS spectra of the magenta ink samples, Figs. 11 and 12, confirmed the dominant content of calcium in magenta ink formulation. Proportions of all inorganic elements in magenta ink samples, given in Tables 4 and 5, indicated that the amount of calcium decreased in magenta ink sample 2 because the higher interaction between magenta ink and fountain solution causing absorption of calcium ions into the fountain solution. This is the reason why pH, conductivity, total dissolved solids and Ca²⁺ ion concentrations in fountain solution dramatically increases in sample 7.

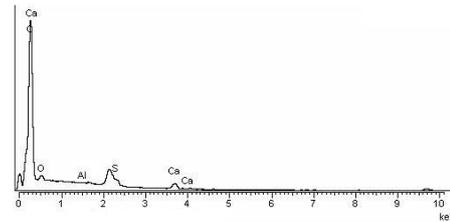


Fig. 11 EDS elemental spectra of reference magenta ink sample (sample 1)

Table 5 Chemical composition of reference magenta ink sample (sample 1)

Element	mass %
C	88.66
O	9.49
Al	0.09
S	0.84
Ca	0.92

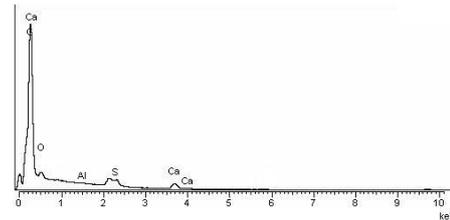


Fig. 12 EDS elemental spectra of magenta ink sample 2, when 12,000 sheet was printed

Table 6 Chemical composition of magenta ink sample 2

Element	mass %
C	85.68
O	12.35
Al	0.09
S	0.80
Ca	0.87

3.3 Image quality assessment

The investigation have shown that changes in the fountain solution quality during the printing process had no influence on the quality of tone, dot, line and text of cyan, yellow and black process color, therefore, the paper represents only the results of influence the examined physical-chemical parameters of fountain solution on the quality of magenta prints and their mutual correlations.

3.3.1 Tone quality

Although, there are many variables that influence tone value increase (TVI) it was of great importance for this investigation to examine how the changes in the fountain solutions quality during the printing of 24,000 sheets impact on the mechanical and optical tone value increase of magenta color. The changes of TVI of magenta color during the printing 24,000 sheets are presented in Fig. 13a-m. The measured TVI on the printed sheets (1-13), Fig. 13a-m, indicated the significant differences in the mid tones of magenta color. The TVI of samples 1-6 is gradually

increased with increasing pH, conductivity and Ca^{2+} ion concentrations. In samples 7 and 8 TVIs were more than 40% for 50% tone value. The possible reasons for that are: the quality of fountain solution, improperly water/ink balance, incompatible the fountain solution component and magenta ink and color order during the printing (MCKY). As can be seen from Figs. 13g and 13h (samples 7 and 8) optical TVI was the same as mechanical. After sample 8 TVI began to decrease by following the trend of the examined physico-chemical parameters up to sample 13 when the value of TVI again increased.

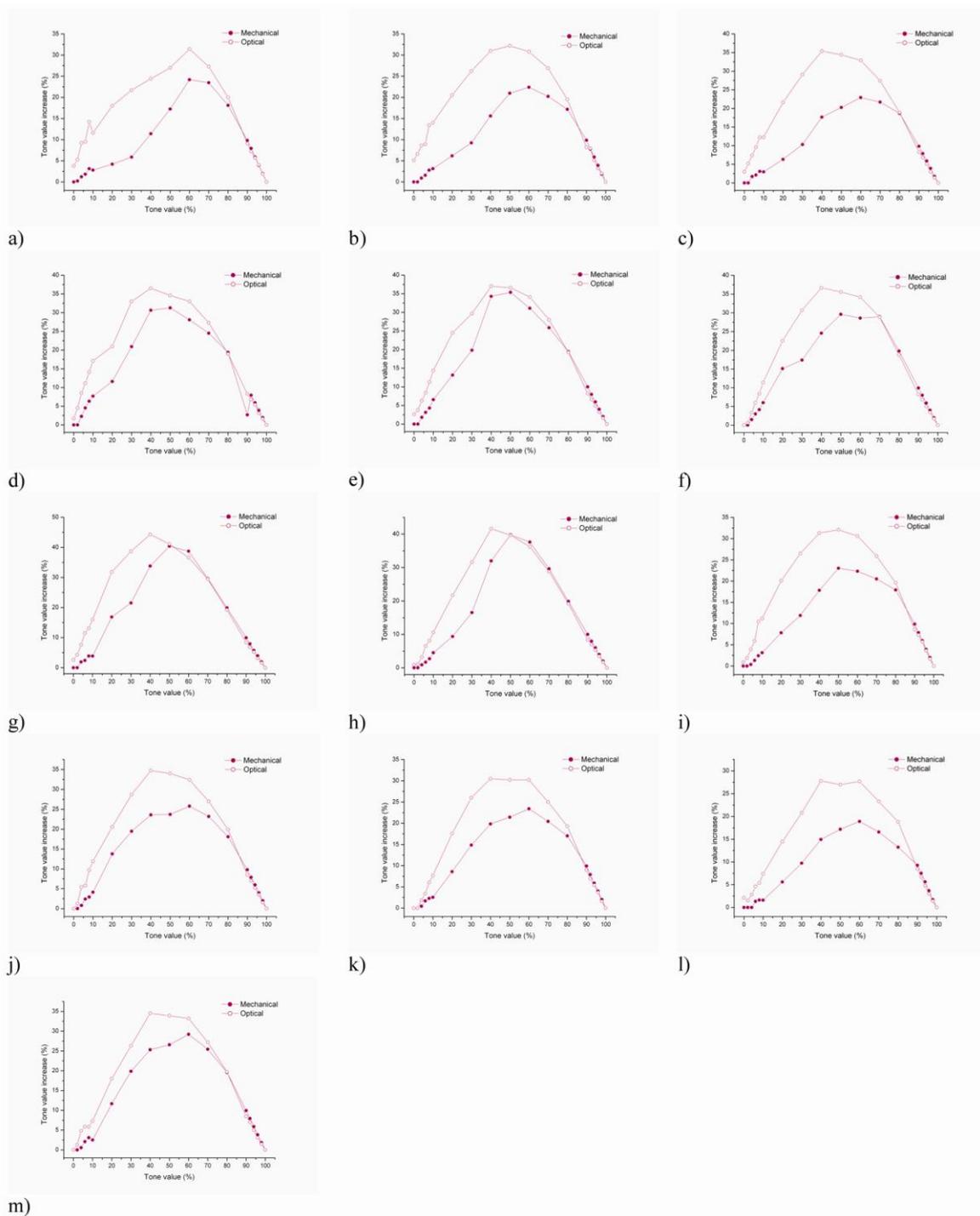


Fig. 13 Mechanical and optical tone value increase of magenta color during the printing of 24,000 sheets

By comparing the measured values of the physico-chemical parameters of fountain solution (Figs. 5 and 6) with the measured tone value increase (Fig. 13a-m) it can be concluded that the changes of pH value, conductivity and Ca^{2+} ion concentrations in fountain solution samples had the significant influence on the formation of magenta dots. The reason for the observed correlation is greater absorption of calcium ions from the magenta ink in the fountain solution during the printing.

3.3.2 Dot quality

One important factor for print quality is the sharpness and contrast of halftone dots that are used to print continuous-tone images. The sizes of the dots on the prints must not be significantly modified, i.e. they must not change either the size or their geometrical shape (circularity).

The results obtained in ImageJ software showed the deviations from the ideal value of circularity (value 1) of magenta dots during the printing of 24,000 sheets (Fig. 14). Magenta dots were ragged and had irregular, distort shape with the least circularity value of 0.36. Also, the significant decreasing and increasing of magenta dot circularity was observed between the analyzed printed sheets.

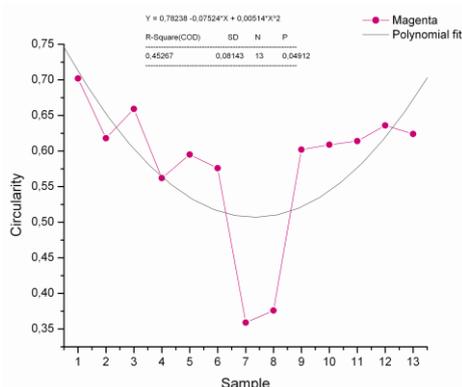


Fig. 14 Dependence of magenta dot circularity on the printed sheets (1-13)

As can be seen from Fig. 14 between magenta dot circularity and samples of printed sheets exist the nonlinear dependence, described by second-order polynomial with the equation curve: $y = 0.00514x^2 - 0.07524x + 0.78238$. This correlation is weak due to the correlation coefficient $R^2 = 0.4527$ and standard deviation $SD = 0.08$.

Comparison of the curves from Figs. 5, 6 and 14 suggested that magenta dot circularity is in the correlation with pH, conductivity and Ca^{2+} ion concentrations during the printing of 24,000 sheets, i.e. increasing of pH, conductivity and Ca^{2+} ion concentrations in fountain solutions caused decreasing of magenta dot circularity, and vice versa. The reason for those correlations is the presence of calcium ions in the contaminated fountain solution and

changes of their concentration in magenta ink during the printing. Generally, the results of magenta dot circularity indicated that the dot reproduction was not consistent during the printing of 24,000 sheets due to the high standard deviation (up to 0.08) and the dissimilarity of dot shape.

3.3.3 Line and text quality

It is known that an image quality must not be judged based on a single entity such as dot circularity, but the reproduced line and text, as the important elements of each image, must also be considered. Line and text analysis is easier since line and text primarily depend on the value of 'raggedness' (the geometric distortion of the edges of the line and text). Since ISO 13660 standards [11] defines only the minimum length of line (1.25 mm) this investigation included the analysis of horizontal and vertical lines that were longer than the minimum defined length. A good quality line and text is described as the one having the least raggedness and sharp edges. The value of line raggedness is expressed in the function of the area ($P = ab$) and the perimeter ($O = 2(a+b)$). The results of line area (Figs. 15a and 16a) indicated that vertical 1.5pt line printed with magenta color (with 1.9 to 5.4% area increase) and horizontal 1.5pt line printed with magenta color (with 6.4 to 10.5% area increase) had the most raggedly edges.

Line raggedness as a function of line perimeter, Figs. 15b and 16b, showed also that vertical 1.5pt line printed with magenta color (with 2.2 to 8.2% perimeter increase) and horizontal 1.5pt line printed with magenta color (with 1.0 to 7.5% perimeter increase) had the most raggedly edges. In the both cases, the least raggedly edges are obtained on the tiny horizontal and vertical lines (1pt). The results indicated that the area and perimeter of analyzed lines increased during the printing due to the ink smearing on the line edges (Figs. 15 and 16).

All the prints with a text value up to 12 units showed no eye visible difference when they was visually checked, but by scanning with the resolution of 1200 dpi text look raggedness.

For the serif font, Fig. 17a, the least raggedness is observed with 10pt font size (up to 7.5% area increase); whereas 5pt font size had the most raggedly edges (up to 41.2% area increase). The sans-serif font (Fig. 17b) showed significantly better results in comparison with the serif font. The least raggedness is observed with 9pt font size (up to 5.4% area increase); whereas 4pt font size had the most raggedly edges (up to 12.3% area increase). The text analysis also showed that the text area increased during the printing due to the ink smearing on the text edges (Fig. 17a and b).

The analysis showed that there was no mutual dependence between the physico-chemical parameters of fountain solution (Figs. 5 and 6) and the line and text raggedness (Figs. 15-17).

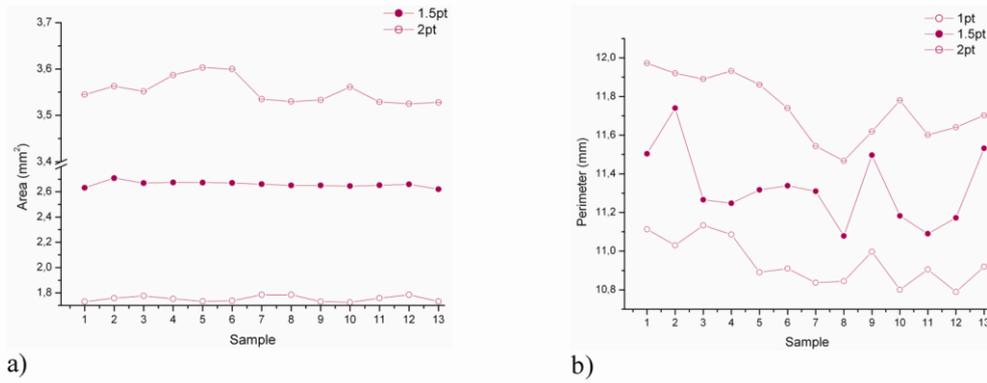


Fig. 15 Vertical line raggedness as a function of: a) line area and b) line perimeter

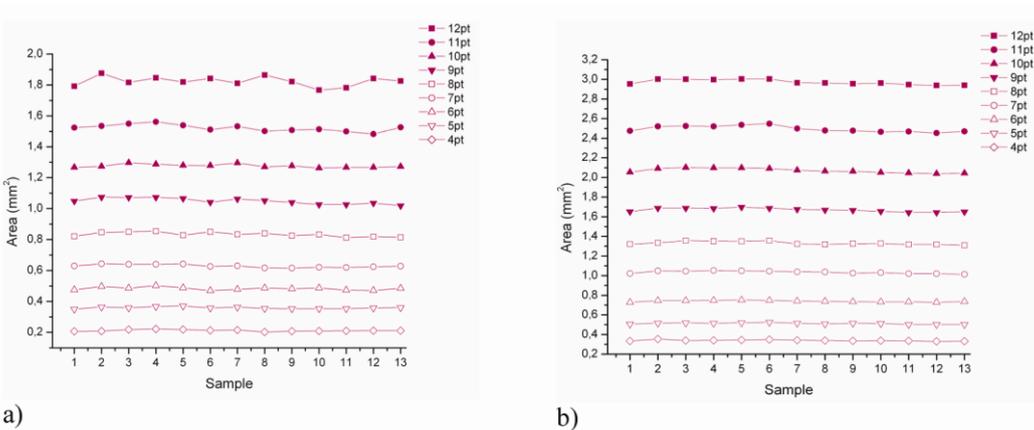


Fig. 16 Horizontal line raggedness as a function of: a) line area and b) line perimeter

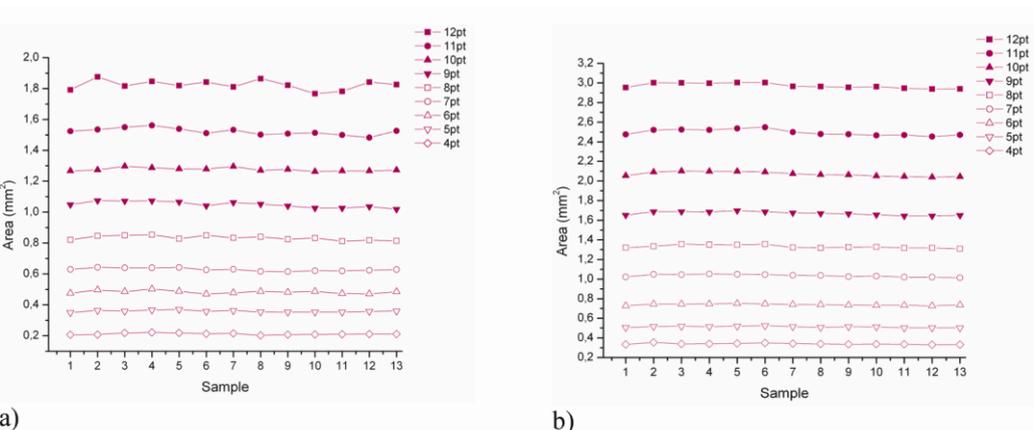


Fig. 17 Text raggedness: a) Times and b) Arial font

IV. CONCLUSION

The experimental data of the influence of physico-chemical parameters of fountain solution on the print quality indicated that the values of pH, conductivity and Ca²⁺ ion concentrations of fountain solution significantly varied during the printing 24,000 sheets with the increasing tendency of all parameters in sample 7, when 12,000 sheets was printed. The changes of pH value, conductivity and Ca²⁺ ion concentrations during the printing were confirmed by the EDS characterization of used printing materials (paper and magenta ink).

The highest mechanical and optical tone value increases were observed on samples 7 and 8 (when 12,000 and 14,000 sheets were printed) due to the changes in fountain solution quality, improperly water/ink balance,

incompatible fountain solution component and magenta ink and color order (CMYK) during the printing. The measurements in ImageJ software showed the considerable deviation of magenta dot circularity. Between the magenta dot circularity and the printed sheets was existed the mutual nonlinear dependence, described by second-order polynomial with equation: $y = 0.00514x^2 - 0.07524x + 0.78238$. The analysis of geometric distortion of the line and text printed with magenta color showed that 1.5pt horizontal and vertical lines had the most raggedly edges, whereas the least raggedness was obtained with 1pt horizontal and vertical lines. The sans-serif font, Arial, showed the least raggedly edges in all range of font size (4-12pt), in comparison with Times.

The analysis shown that between the lines and text printed with magenta color and pH, conductivity and Ca^{2+} ion concentrations of fountain solution was not existed mutual dependence. However, the changes in pH, conductivity and Ca^{2+} ion concentrations had the significant impact on the quality of tone and circularity of magenta dots.

REFERENCES

- [1] Brigl & Bergmeister GmbH (B&B), Biogloss, [Online] www.brigl-bergmeister.com, Accessed 25th May 2012.
- [2] W. Chen, X. Tang, J. Considine, K. T. Turner, Effects of inorganic fillers in paper on the adhesion of pressure-sensitive adhesives, *Journal of Adhesion Science and Technology*, 25, 2011, 581-596.
- [3] S. S. Deshpande, Fountain solution in lithographic offset printing, *Journal of Engineering Research and Studies*, 2(2), 2011, 82-88.
- [4] A. Dhopade, Image quality assessment according to ISO 13660 and ISO 19751, *Image Quality Assessment*, 2011, 43-50.
- [5] P. D. III Fleming, *Printing, Pulp and Paper*, 2004, 1-9.
- [6] J. C. Froberg, J. Voltaire, M. Sundin and F. Tiberg, Effect of ink-fountain balance on ink-tack development, *2000 TAPPI International Printing and Graphic Arts Conference Proceedings*, Atlanta, USA, 2000, 133-138.
- [7] B. Fuchs, The influence of the quality of raw water in damping solutions on newspaper printing, IFRA, Special Report 1.15, 1996, 1-22.
- [8] Huber group, Safety Data Sheet for Inkredible RAPIDA F 10 RP, [Online] <http://www.hubergroup.info/lang/en/everpskala.html>, Accessed 25th May 2012.
- [9] R. Hundal, Calcium Control in the Pressroom. Alkaline Paper Advocate, 10(4), [Online] <http://cool.conservation-us.org/byorg/abbey/ap/ap10/ap10-4/ap10-402.html>, Accessed 23 May 2012.
- [10] Ink & Print, Calcium and it Effects on the Printing Process, [Online] <http://www.inkandprint.com/Tech-Tips/calcium-and-it-effects-on-the-printing-process.html>, Accessed 20 May 2012.
- [11] International Organization for Standardization (ISO), ISO12647-2:2004 Graphic technology - Process control for the production of half-tone colour separations, proof and production prints - Part 2: Offset lithographic processes, 2004.
- [12] International Organization for Standardization and International Electrotechnical Commission (ISO/IEC), ISO/IEC 13660:2001(E) Information Technology - Office Equipment - Measurement of image quality attributes - Binary Monochrome text and graphic images, First edition, 2001.
- [13] International Organization for Standardization (ISO), ISO/WD 19751-1 Office Equipment - Appearance-based image quality standards for printers - Part 1: Overview, procedure and common methods, A Working Draft of Technical Committee ISO/IEC JTC1 SC 28, 2004.
- [14] E. E. Kiš, A. G. Lomić, R. P. Marinković-Nedučin, *Eksperimentalna fizička hemija* (Tehnološki fakultet, Novi Sad, 1998). (In Serbian)
- [15] H. R. Leach, J. R. Pierce, *The Printing Ink Manual* (Springer: Netherlands, 1993).
- [16] A. P. Lewis, Colored Organic Pigment, in C. D. Craver and C. E. Jr Carraher (Ed.), *Applied Polymer Science: 21st Century*, (Elsevier Science: UK, 2000) 493-526.
- [17] M. Mahajan, Fountains: not so difficult as we may think, *Panpa Bulletin*, 2006, 38-39.
- [18] K. Rizzo, Fountain Solutions - The Details. Printing Industries of America/Graphic Arts Technical Foundation (PIA/GATF), *GATFWorld*, 2007, 16-18.
- [19] D. R. Rothbard, Electron Microscopy for the Pulp and Paper Industry, in Z. R. Li (Ed.) *Industrial Applications of Electron Microscopy*, (New York: Marcel Dekker, Inc., 2003) 33-49.
- [20] Sappi Europe SA, Paper, Ink and Press Chemistry. Exploring key print variables, 2004, 1-21.

Efficient Utilisation of C.B.M in Industry for Reducing Cost of Maintenance

Amit Kumar Jain¹, Rahul Maurya², Vineet Kumar Singh³

¹ Research Scholar, M-Tech (Maintenance Engineering & Management), Department of Mechanical Engineering M.A.N.I.T. Bhopal, INDIA

² M-Tech (Maintenance Engineering & Management) MANIT Bhopal India.

³ M-Tech (Hydro Power Engineering) MANIT Bhopal India.

Abstract: Every industry is dependent on some type of asset that keeps the business in business be it a computer, a CNC, or a megawatt transformer. In a large enterprise, reducing costs related to asset maintenance, repair, and ultimate replacement is at the top of management concerns. Downtime in manufacturing ultimately results not only in high repair costs, but in customer dissatisfaction and lower potential sales. In response to these concerns, this paper presents a methodology for creating an efficient Condition Based Maintenance (CBM) system in industry for reducing cost of maintenance. It provides valuable guidelines for planning an enterprise system that monitors critical maintenance processes and assets.

Keywords: CBM, Downtime, Maintenance, Condition Monitoring, CBM+,

I. Introduction

The machines/equipments of every industry are his most valued assets. These are his source of income and thus they should be accorded the care that they deserve. Condition based maintenance generally is appropriate for machines that are highly relied on in every organization.

Condition Based Maintenance is a predictive maintenance technique. CBM is the method adopted to monitor and diagnose the condition of the process, machinery, or components under investigation. In short we can say that CBM is a technique of diagnosing failure mechanisms and making a prognosis of the remaining useful life before failure. Efficient maintenance systems are expected to detect early forms of degradation in predictive maintenance (pdm) practices in tandem with Condition Based Maintenance (CBM).

Condition based maintenance is triggered by a predefined event which indicates the extent of degradation of component or equipment. The difference between time based (periodic) and condition based maintenance can be defined in terms of the way in which they aim to prevent functional failure [1]:

- Time based (periodic) maintenance entails performing an action without first collecting component/equipment data;
- Condition based maintenance involves, on the contrary, collecting data, conducting an analysis, and performing a maintenance action if necessary.

Condition based maintenance leads not only to a reduction of unnecessary preventive maintenance actions and in shifting some maintenance operations away from the outage period (on-line maintenance) but ensures that actions are performed when justified by component/equipment condition. Condition based maintenance also helps ensure closer relation between operations and maintenance departments as well as among different maintenance personnel of different background skills.

CBM is carried out for two main reasons:

1. To detect sudden changes in condition that could lead to catastrophic failure, particularly for machinery that could represent a threat to the health and safety of people, or cause an environmental incident. This is known as 'Machinery Protection' or 'Protective Monitoring'.
2. To identify the early onset of incipient failures so that a prediction can be made about their most likely progress and suitable actions can be planned. This is known as 'Predictive Monitoring' or 'Predictive Maintenance', and is often abbreviated to 'PDM'.

An efficient CBM program is highly valuable for every industry in terms of productivity, quality, inventory control, and expenditure on plant and machinery; it also reduces the cost involved in false maintenance which leads towards increased profit for the industry.

II. EFFICIENT CBM

The term efficient implies that a CBM system is capable of understanding and making decisions without human intervention. Technologies making this possible include: sensors with built in intelligence (SMART Sensors) capable of transmitting relatively rich, high grade information; re-programmable on-line sensors, designed to be reconfigured with new rules in the event that detectable recognisable patterns change; algorithms, fuzzy logic and neural networking, designed to analyse trends within recovered sensory data, and produce decisions on the likelihood of failure of monitored plant items [2]; artificial intelligence algorithms capable of providing proxy data as a substitute for failing or a failed sensor, whilst the malfunctioning sensor is repaired [3]. Further it is also possible through integration of a CBM system with a company's computerised purchasing system, thus automating parts ordering [4].

As technological advancements have fed into CBM so the method of deploying CBM systems and integrating them with other business systems has changed.

It takes place gradually as new scientific discoveries are made, accepted and applied to CBM systems. Recent technological advances include improved knowledge of material failure mechanisms, advancements in failure forecasting techniques, advancements in monitoring and sensor devices, advancements in diagnostic and prognostic software, acceptance of communication protocols, developments in maintenance software applications and computer networking technologies to name a few.

III. IMPORTANCE OF CONDITION MONITORING

Condition monitoring involves determining the condition of a machine and its rate of change of measured parameters in order to determine the maintenance requirement. The condition of machine may be determined continuously or at regular intervals by monitoring measurable parameters.

The most common of types of techniques used in condition monitoring today include human senses which are visual observations, listening and touching. Such senses employ visual inspection, tactile inspection smelling, aural inspection and optical magnification [5]. An efficient condition based maintenance program shall utilize variety of technologies, the majority of industry equipment constitute of mechanical systems. An efficient condition based maintenance program must include various monitoring and diagnostic techniques. These techniques include:

- Vibration monitoring.
- Acoustic analysis.
- Motor current analysis technique.
- Motor operated valve testing.
- Thermography.
- Tribology.
- Process parameter monitoring.
- Visual inspections.
- Other non-destructive testing techniques.

A pre-requisite for condition based maintenance is the availability or development of non-destructive monitoring techniques. The goal is to detect degraded condition, and if possible, assess changes in them with a view to preventing failure.

IV. EXECUTION OF EFFICIENT CBM

The execution of efficient condition based maintenance can be done in the following four steps:

1. *Data collection.* The relevant data is collected (off line or online) through the use of process control systems, vibration measurements, oil sampling, and other methods.
2. *Data analysis.* Depending on the situation, the data needs to be cleaned; for example, during start-ups and shutdowns the plant may exhibit erratic behaviour, which is not to be misinterpreted as failure. The data can be analysed in several ways, for example by direct comparison with a threshold or by looking at trends or other remarkable behaviour.
3. *Decision making.* Based on the data and the analysis, a decision is made. Such a decision may involve a change in operating routines or the direct execution of a maintenance task. It may also lead to additional data collection and analysis.
4. *Implementation.* When a decision has been made, an intervention is planned. After the intervention, reports can be made and stored for future maintenance actions. Evaluations are conducted when deemed necessary.

V. BENEFITS ACHIEVED FROM EFFICIENT UTILIZATION OF CBM

- Reduction of scheduled preventive maintenance content.
- Escalation of preventive maintenance intervals.
- Early warning of impending failures through data link.
- Enable scheduling of unscheduled maintenance.
- Enhanced diagnostics for troubleshooting.
- Optimized logistics support.
- Pre-staging replacement parts for impending failures.
- Fewer components replaced for preventive maintenance over the life of the equipment.

The benefits of CBM can be enhanced if industries implement CBM+ it will also greatly support the maintenance system of the industry.

Conditioned Based Maintenance Plus (CBM+) is the application and integration of appropriate processes, technologies, and knowledge-based capabilities to improve the reliability and maintenance effectiveness of industrial systems and components. CBM+ is maintenance performed on evidence of need provided by Reliability Centered Maintenance (RCM) [6] analysis and other enabling processes and technologies. CBM+ uses a systems engineering approach to collect data, enable analysis, and support the decision-making processes for system acquisition, sustainment, and operations.

CBM+ encourages the use of portable maintenance aids to:

- Reference technical manuals and checklist
- Assist in production management
- Enable data recording and recall
- Provide a platform for training
- Access embedded diagnostics and prognostics

VI. Conclusion

In this paper we present an efficient condition-based maintenance approach for industries, which will improve maintenance performance and machine utilization. And the following key benefits can be achieved:

- Improved system reliability.
- Decreased maintenance costs.
- Decreased number of maintenance operations causes decreasing of human error influence.
- Increased profit.

VII. Acknowledgement

We are thankful to Department of Mechanical Engineering of MANIT Bhopal India for providing the required facilities needed for the successful completion of this paper.

References

- [1] Proceedings of the IAEA Specialists Meeting on Nuclear Power Plant Condition Monitoring and Maintenance, Lyon, France, 2-5 June 1998.
- [2] Philip Jacob, 1994, "[Ideas for the exploitation of distributed intelligence in condition monitoring systems](#)", Maintenance Engineering Research, The University of Manchester.
- [3] Rolf Orsagh, Mike Roemer, and Ben Atkinson, 2000, "[An internet-based machinery health monitoring](#)."
- [4] J.Coy, R.M. Parkin, M.R.Jackson, N.Steward, 2001, Proceedings, Int Conf Mail Technology; Evolution to e-Revolution, 24-25 April, Hilton Brighton Metropole, UK, ISBN 186058 327X, 99 383-391.
- [5] Implementation Strategies and Tools for Condition Based Maintenance at Nuclear Power Plants IAEA TECDOC 1551 (ISBN: 92-0-103907-7) Date Published: 2007.
- [6] Krishna B. Misra, Handbook of Performability Engineering, Pp1316 Springer 2008 ISBN 978-1-84800-1305.

Preliminary Model Development of Client Orientation in Civil Engineering Businesses

Christopher Nigel Preece, Hooman Abadi

Department of Civil Engineering, Razak School of Engineering & Advanced Technology, University Teknologi Malaysia (UTM), Malaysia

Faculty of Civil Engineering, University Teknologi Malaysia (UTM), Malaysia

ABSTRACT: As a key pillar underlining the marketing concept, client orientation is a philosophy in that an organization's customers should be put in the centre of its activities; therefore, it is able to create superior value for them continuously. Competing in the rapidly changing and highly competitive marketplace, firms in civil engineering sector need to attain this. This paper analyses various elements which contribute to the achievement of client orientation in civil engineering by starting with a review of the competitive marketplace and the approaches that the civil engineering sector have been adopted to meet the new challenges, then definition of client orientation. In conclusion, the preliminary "C-O-N-S-T-R-U-C-T-S" framework is proposed which aims to provide an outline approach to achieving client orientation in the civil engineering sector.

Keywords: Client Orientation, Construction, Marketing, Civil Engineering

I. INTRODUCTION

Clients in the construction industry comprise of a variety of organizations different in size, sector and geographical location, culture etc. They have been seen as the "driving force" of the industry, and therefore their needs and wants must be satisfied (Latham, 1994). Clients demand that products and services are delivered to time, budget and quality, work from day of handover, add value, reduce business risk (Williams, 2001). However, there have been rapidly significant changes in clients' expectations recently, which have posed challenges on the construction industry (Blockley and Godfrey, 2000), as well as on the Civil Engineering Sector. Clients' expectations, and therefore, their requirements to the market have been driven to change according to the issues both from their own, and from the external environments. There is a continuing and increasing demand of clients for good quality and performance improvement on construction products and services (CITB, 2004; Blockley and Godfrey, 2000). Innovation and new

Technologies that the industry is being encouraged to apply have had more of an impact on the firms, requiring them to be more receptive and ready to change. The new requirements from legislation, for example policies in work-life balance, health and safety issues or in reducing the damage to the environment, etc... are considered as significant and influential factors to clients' needs and wants. In particular, the Government's policy for sustainable development reinforced by legislation has impacted the industry as well as on clients' requirements (CITB, 2004). In addition, the world is becoming less predictable, necessitating greater flexibility, and increased globalization

means more international clients and also more competitors. Those challenges have made the market more and more competitive. The changes in clients' profiles, which is illustrated by the fact that the workload that the sector carried out for public sector clients has decreased while that for the private sector has been increasing (Latham, 1994), have also influenced the market.

Therefore, whether tendering for projects in the domestic construction market or overseas, Civil Engineering businesses have to face the fact that their clients now have not only a greater range of requirements but also of choice than ever before.

II. Approaches of the Civil Engineering Sector to meet the Changing Environments

In response to the rapid changes from environments, the Civil Engineering sector, together with the whole construction industry, has been adopting new approaches to face the new challenges from the market.

Latham (1994), cited by Blockley and Godfrey, (2000) recognized the interdependence of the stakeholders in construction, such as clients, design professionals, contractors and subcontractors, and therefore articulated the need to construct the team inclusive of them.

Supporting the idea of Latham, in his report Rethinking Construction, Egan (1998, cited by Blockley and Godfrey, 2000) pointed out that construction industry, including the Civil Engineering Sector, "must replace competitive tendering with long-term relationships". Egan also indicated "five key drivers of change" for the industry: committed leadership, focus on the customer, integrated processes and

Teams; a quality driven agenda, and commitment to people. Therefore, partnering, in which the relationship among participants plays the most important role, becomes essential to construction industry's survival and growth (Critchlow, 1998). According to Latham (1994), some clients "have taken steps to enhance the level of consumer satisfaction for themselves and their tenants/occupiers". The Highways Agency even makes use of the Capability Assessment Toolkit (CAT) to help establishing and maintaining the relationship within its supply chain management (Highways.gov, 2006). Other tools and techniques applied may include the EFQM Excellence Model (catalystconsulting.co, 2006), Constructing Excellence (constructingexcellence.org, 2006), etc.

Civil Engineering businesses, a group of "players" in construction, have realized that they "need to focus on delivering value and to delighting the client and his customers through a much deeper understanding of their wants and needs from the value chain" (Blockley and Godfrey, 2000). This is a significant change from the traditional marketing approach in which the focus of Civil

Engineering firms is the project, not the customer (Cova, Ghauri and Salle, 2002). They now admit that “one of the most powerful” approaches in order to win projects is “to make the current customer happy” (Courtis, 1987).

They can win customers and outperform competitors by doing a better job of meeting and satisfying customers’ needs (Preece et. al., 2003).

Therefore, as a key pillar underlining the marketing concept, client orientation is a philosophy that puts customers and clients at the heart of business activities. However, a search of the literature shows that little attention has been afforded this by civil engineering businesses or the construction sector in general, though the client orientation philosophy, programmes, tools and techniques are well established in business generally. The problems discussed in the next section may be the obstacles that prevent organisations in Civil Engineering from implementing this philosophy.

III. Client Orientation and the Issues in Implementing It in Civil Engineering

Christopher et. al (1991) have distinguished six markets in which stakeholders may be regarded as clients and suppliers in some dimensions. They include: customer markets, internal markets, referral markets, suppliers markets, potential employee markets and influence markets. However, this paper focuses only on the customer markets, where Civil Engineering businesses play as suppliers of services to their clients, who commission the project or activities and enter into the contractual relationships with them, as the definition suggested by Pettinger (1998).

Client orientation involves taking actions based on market intelligence, not on verbalized client opinions alone (Kohli & Jaworski, 1990). It requires the organization to stay close to its customers and track their satisfaction over time (Preece et. al, 2001). It is the sufficient understanding of one’s client to be able to create superior value for them continuously, including customer commitment; creating customer value; understanding customer needs; identifying customer satisfaction objectives; measuring customer satisfaction; and providing after-care services (Narver & Slater, 1990). To achieve this philosophy, the organization needs to obtain and use information from their client, develop a strategy which will meet client needs, and implement that strategy by being responsive to clients’ needs and wants.

It is therefore necessary to identify the issues that may frustrate development of client orientation in the construction industry. Preece ET. al (2003) conclude 5 issues as following:

Issue 1: Management of marketing, quality and the human resource are relatively underdeveloped areas in the construction industry. Emphasis has therefore been on technical details such as technology innovation and the quality of the end product such as the building or structure, rather than on developing services to exceed the needs and wants of the clients and achieving their satisfaction.

Issue 2: The standard of service in providing construction projects is often affected by major conflicts and disputes between parties, rather than a partnership culture and collaborative or alliance strategy focusing on satisfying the clients’ needs and wants.

Issue 3: The people in contract with the client in construction change during the course of a project.

This places great emphasis on all a firms personnel being trained to be ‘caring’ to the client team. In addition, subcontracting and outsourcing of responsibilities, make it more difficult to control the way the client team is being served.

Issue 4: Clients’ perceptions are ‘coloured’ by the overall image of the industry.

Issue 5: Given the low profit margin in contracting, client orientation programme is likely to be viewed as too expensive to implement.

In addition, there are some other issues connected with the support from the top management, appropriate delegation or empowerment, effective communication and the recognition in the organization, are also possible barriers to achieving client orientation.

However, except for the increasing awareness of Civil Engineering firms, there are some favorable conditions to apply client orientation into the sector. First, the nature of the construction industry has led to the unique relationship between client and constructor. Clients and constructors have to be involved with high degree of uncertainty in a long transaction with many phases clearly delimited. Therefore, both sides in the relationship naturally recognize the need to keep each other informed about the projects. In addition, although clients vary, an increasing number of them are “well informed, know what they want and take decisive steps to achieve it” (Latham, 1994). The orientations to constructing teams and partnering in construction industry also advocate the application. Last but not least, the support from the Government, the biggest client of the sector, could be considered as another advocacy.

IV. The preliminary development of the C-O-N-S-T-R-U-C-T-S framework

At present, client orientation in the construction industry is something that needs further development and has not been fully established and considered by many businesses.

The main ideas behind client orientation are still not fully understood even through the client orientation philosophy is crucial to the success of the construction company in highly competitive markets.

Confronted with the uniqueness and complexity of each project as well as the issues faced by implementing client orientation in the construction industry, a “C-O-N-S-T-R-U-C-T-S” framework of client orientation has been developed. This is based upon a general framework proposed by Cook (2002) and has been piloted with a small number of construction businesses. In her research, Cook suggested 9 elements that “need to be in place” to help firm to create a climate and culture of customer orientation. They include:

1. Business imperative and top team clarity,
2. Listening posts,
3. Service strategy and goals,
4. Customer-driven processes,
5. People and development,
6. Empowerment,
7. Communication,
8. Reward and recognition, and

9. Sustaining a customer focus.

Considering the general management perspective and its mitigation of the issues that may frustrate the development of customer orientation in Civil Engineering Sector, the proposed C-O-N-S-T-R-U-C-T-S framework has been developed, as in Table 1. Each of its elements is given a weighting score, determined by using a set of questions to gather information from relevant sources in the firm, for the measurement of Customer Orientation level as in Table 2

Table1. C-O-N-S-T-R-U-C-T-S framework

Factors	Description	Weighting score	Frustrated issues related
1. C	Clarity of the top management team: The growing importance of client orientation as a mean of differentiation in the construction industry must be recognized by senior managers and project managers, as it is they who fashion the response of their organizations to client orientation. It is from the top management that a client orientation programme could be started. Through the top management team, the mission, vision, values and strategy related to client orientation could be established.	10	1
2. O	Operate the current situation analysis: The construction company attempts, usually through market research to conduct the customer service audit operation—internally and externally. From the external, the relative importance of time, cost and quality must be given careful consideration and be discussed with the client. . It is also necessary to ask employees especially project staff what they consider would be unsettling for the customer in the transaction of the project with their department.	5	/
3. N	Need of value management and risk: Value management is chosen to make the functional benefits of a project more explicit consistent with a value system defined by client (Kelly, Male and Graham, 2004). In addition, risk management, which is used to identify the possible risks and use appropriate methods to mitigate the negative effects, together with the value management, aim at the achievement of the customer’s best value for money.	7	5
4. S	Supply chain management : Supply chain management is a strategic function of the construction form that integrates those external and internal activities required to manage the sourcing, acquisition and logistics	6	2

Factors	Description	Weighting score	Frustrated issues related
	of construction resources essential for the organization to produce construction project or services that add value to the customers (Smith, 2002).		
5. T	Training and people development: Training and development is an essential cornerstone in promoting a customer service philosophy (Cook, 2002). It is crucial to include everyone throughout the construction company especially the project team in training and development to enhance their knowledge, skills and attitudes towards client orientation.	9	1, 3
6. R	Relationships with the external clients Developing long-term relationships with loyal customers and retaining existing customers could help to get valuable customer information and reach win-win situation.	6	4
7. U	Understanding, commitment and satisfaction by employees The attitude taken by staff is related to the successful achievement of the customer orientation and customer care.	6	3
8. C	Communications A well-defined system of communication is essential to enable rapid dissemination and feedback of management and staff views, proposals and actions in the implementation of the customer orientation philosophy.	8	4
9. T	Tests on motivation for client orientation Test results through customer satisfaction surveys, evaluate the training methods. Create a motivating climate for their employees by recognizing and rewarding client-oriented behavior, and then improve the client orientation programme on a continuous basis.	5	/
10. S	Services of after-care As in any other kind of service organisation, the quality of after-care customer service is a critical factor in the winning of repeat business, and it is not exception in the construction industry.	8	/

Table2. Standards for the Measurement of Customer Orientation

Scores	Level of Customer Orientation
49 +	★★★★★
42 ~ 48	★★★★
35 ~ 41	★★★
28~ 34	★★
27 or below	★

In order to demonstrate the effectiveness of the framework, a pilot survey has been conducted. In this research, a questionnaire designed for face-to-face interviews, using a qualitative approach has been used to create data related to each element in the proposed framework.

The following conclusions have been drawn out from the information gathered:

1. It is believed that customer orientation and the understanding of it are critical to operate business profitably and create competitive advantage. This conclusion supports all of the elements in the framework, particularly elements 1, 5, and 7.
2. Supply chain management in construction industry should be given more priority; therefore it should be a separate element as in the framework, as a development point from the Cook's framework.
3. Some of the respondents agree that those elements are included in their policy to implement customer orientation, such as the Balanced Business Scholar-Card or KPI.
4. Most of the respondents agree that the elements in the framework are critical for a successful application of customer orientation.

In line with their suggestions, the framework would seem to require the following improvements:

- I. Cancel the information about risk management
- II. Add some questions about the target companies sizes and market share
- III. Consider the relationship among the main contractor, architect, consultant and subcontractor.
- IV. Improve the set of questions enclosed.

The framework, after those improvements, is illustrated in Figure 1.

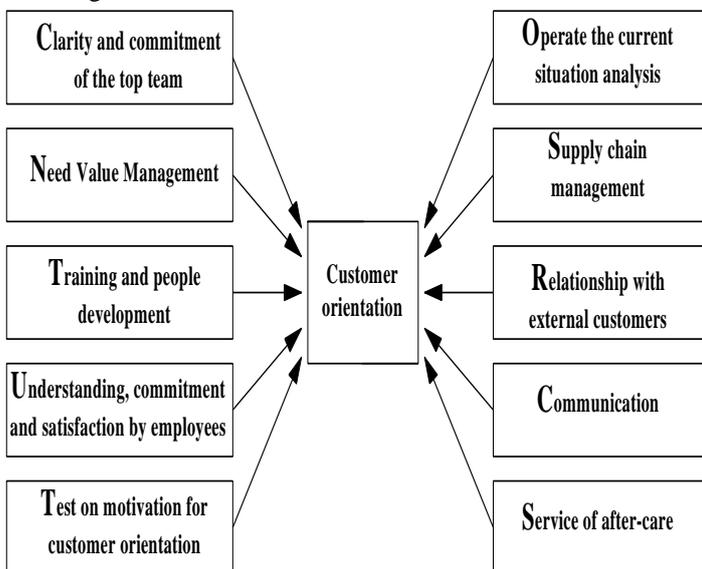


Figure 1: “C-O-N-S-T-R-U-C-T-S” Framework of Client Orientation in Construction Industry

V. CONCLUDING & COMMENTS

The expectations and requirements of clients for construction industry in the UK, the “driving force” of the industry, have rapidly been changing recently. This fact has put pressure on the way Civil Engineering businesses operate. New approaches have been adopted to help them to confront to the new challenges. After Latham's idea

of “constructing the team” among stakeholders in construction, Egan (1998) emphasizes the need of firms to build a long-term relationship with their clients. Partnering becomes essential to construction industry's survival and growth (Critchlow, 1998, p.2). Many other tools and techniques such as CAT, Business Excellence Model, Constructing Excellence, etc... have been used. Civil Engineering businesses, therefore have realized that they need a much deeper understanding of clients' wants and needs from the value chain. According to Preece and Moodley (2003), there are five key issues that frustrate the implementation of this philosophy in the sector.

Based upon a general framework proposed by Cook (2002) considering the nature of the sector as well as the issues faced by implementing client orientation in the construction industry, a “C-O-N-S-T-R-U-C-T-S” framework of client orientation has been developed, together with a set of questions to help measure Customer Orientation Level in firms. As this study is only a very small part of research into client orientation of construction companies, there is still scope for further investigation into client orientation in the construction industry which can be looked at in the future.

The areas that can be researched further are possibly:

- This research only touched on the proposed framework aimed to achieve client orientation in the construction industry. Therefore, a great deal of time and effort should be invested in commencing the full-scale customer orientation research.
- Comparing and analysing a number of different size or scale construction companies such as small, medium, large and very large size, and try to find how they go about client orientation and differences among these implementation, then, different framework for different size Construction Company needs to be set up.
- Looking at different contractors in different sectors of the construction industry, such as specialised companies, for example, transport focused organisations or water based organisations, how customer orientation could be achieved in professional companies.
- Looking at the international market, the global economy and international construction companies try to find how those international construction companies could meet their foreign customers' needs and wants then achieve client orientation.
- Focusing only on a single construction company which has not been client oriented organisation and study it in depth to find how this company could achieve customer orientation and what the barriers are, then propose appropriate solution methods to deal with these issues.

It is concluded that, this research has just touched onto the topics of customer orientation in the construction industry and there is still a large amount of research need to be conducted in order to understand ideas behind customer orientation in the construction industry.

VI. REFERENCES

- [1] Blockley, D. and Godfrey, P. (2000), "Doing it differently". Thomas Telford Ltd, UK.
- [2] Christopher, M., Payne, A. and Ballantyne, D. (2004), "Relationship Marketing: Creating Stakeholder Value", Elsevier Butterworth-Heinemann, Oxford
- [3] Catalystconsulting.co, (2006)." Website of the Catalyst Consulting Company", URL: <http://www.catalystconsulting.co.uk>
- [4] Constructingexcellence.org, (2006). "Website of Constructing Excellence", URL: <http://constructingexcellence.org>
- [5] CITB (2004) CITB- "Construction Skills", Annual Report and Accounts 2004.
- [6] Cook, S. (2002) Customer Care Excellence: How to Create an Effective Customer Focus. Clays Ltd, UK.
- [7] Curtis, J. (1987). "Marketing Services: A Practical Guide", Kogan Page, London Cove, B., Ghauri, P. N. and Salle, R. (2002), 'Project Marketing: Beyond Competitive Bidding'. John Wiley & Sons, London.
- [8] Crane, A. (2001) "Welcome and introduction". In: Preece, C. (ed.) 2nd International Construction Marketing Conference: Academic Proceedings 2001(cite by Macnamara, 2003)
- [9] Critchlow, J. (1998). "Making partnering work in the construction industry". Chandos Publishing (Oxford) Ltd, UK.
- [10] Egan, J. (1998) "The Egan Report: Rethinking Construction".
- [11] Highways.gov, (2006)."Website of the Highways Agency", URL: <http://www.highways.gov.uk/>
- [12] Kelly, J., Male, S. and Graham, D. (2004)." Value Management of Construction Projects". Blackwell science, Oxford
- [13] Kohli, A.K. & Jaworski, B.J. (1990)." Marketing Orientation: The Construct, Research propositions and managerial Implications", Journal of Marketing, 54 (2), 1-18.
- [14] Latham, M. Sir (1994)." Constructing the team", HMSO.
- [15] Narver, J.C. & Slater, S.F. (1990)."The Effect of a Market Orientation on Business Profitability", Journal of Marketing, 54 (2), 20-35.
- [16] Pettinger, R. (1998)." Construction Marketing: Strategies for Success". Macmillan Press, Basingstoke.
- [17] Preece, C., Moodley, K. and Smith, P. (2003)." Construction Business Development: Meeting New Challenges, Seeking Opportunity". Butterworth-Heinemann, Oxford.
- [18] Smith, K. (2001)." Building relationships, building success". In: Preece, C. (ed.) 2nd International Construction Marketing Conference: Academic Proceedings 2001
- [19] Smith, N.J. (2002)." Engineering Project Management". 2nd edition. Blackwell science, Oxford.
- [20] Williams, A (2001). "The general challenges UK construction marketers face today". In: Preece, C. (ed.) 2nd International Construction Marketing Conference: Academic Proceedings 2001.

Study and Analysis of Dye Sensitized Solar Cells

Thirumal mamidi¹, Shobhan Babu Mandha², Abhishek mishra³

^{1,2} Asst. Professor Dept. Mechanical Engineering M.B.E.Society's Collage Of Engg Ambajogai-431517.

³ Abhishek mishra Dept. Mechanical Engineering VNIT-Nagpur.

Abstract- the need of energy is the basic need for any civilization. Through the years we have been using various kinds of energy sources to fulfill energy requirement. The search for a clean energy source, producing electricity at low production cost is always on. Solar energy is one of the best known energy options available to us. Even if we are able to convert a small fraction of solar energy falling on earth's surface into electricity, then the whole problem related to energy crisis would be resolved. But it has a disadvantage that initial cost for the equipment to harness the energy is very high. That is why it is not very popular in the market. But there is a technology which is making its cost cheaper.

This technology is known as "DSSC". Which stands for, dye sensitized solar cell. DSSC is a new class of solar cell that belongs to the group of thin film solar cells. This solar cell partly by passes the photosynthesis reaction and makes a shortcut conversion of sunlight into electrical current. The inner part of this solar cell consists of semiconductor material. And in between semiconductor molecules, there are dye molecules, similar to chlorophyll in plants. Sunlight is absorbed by them. It makes dye electrons to move to a higher energy level. Then they become mobile and move through semiconductor circuit. This results in the flow of electricity. This cell was developed by Professor "Michael Gratzel".

Keywords: Solar energy, Thin film solar cell, DSSC, TiO₂, Dye, Pyranometer, Overall conversion efficiency, Semiconductor layer thickness and Silicon cell.

I. Introduction

Solar cells have gone through a number of years and they have gone through a number of phases. Their development can be described according to their construction principles. So we can call them solar cells of different generations.

In the first generation of solar cells silicon with a large area and in a single crystal of p-n junction diode was used. This is made from two doped crystals. One is doped with n type material having extra electrons. Other crystal is doped with p type impurities lacking some electrons. When both crystals are placed in contact with each other electrons from n type material starts flowing towards p type material. Eventually enough electrons will flow across the junction to equalize the Fermi level of the two regions. This result in a region at the interface called p-n junction. Where charge carriers are depleted or accumulated on each side of junction. In Silicon the transfer of electrons produces a potential barrier of 0.6 to 0.7V. Right now these solar cells accounts for approximately 86% of the total solar market.

In Silicon solar cells sunlight can provide enough energy to the electrons to move them from valence band to conduction band. This process is known as photo excitation. When an external load is applied these electrons lose their energy while flowing through external circuit. Then they again come back to the n side and recombine with the holes left by them. In this way sunlight creates electrical current. In any semiconductor band gap means the electron with that much energy or higher energy than band gap will contribute to

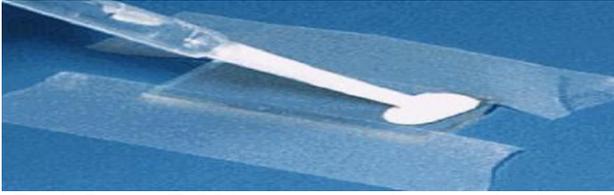
produce electricity. In case of silicon semiconductor they can absorb most of the visible range light from red to violet. But the higher energy electrons at the blue, violet end have the more energy than required so it is wasted. Also there is one more issue that to absorb a good amount of photon energy we have to make n layer thick. This increases material cost and makes the construction complex. This creates one more problem that any electron which is on conduction band can recombine it with a hole on the n region before reaching the p-n junction. This puts an upper limit on the efficiency of Silicon solar cells. However the biggest problem with these cells has been the cost. Silicon processing is costly and the thick layer of Silicon increases its cost more. So we need a different form of solar energy utilization device which would convert solar energy into electricity but at economical price. By the years of observations and research we have got following graph which compares all kind of first generation solar cells efficiencies. We have seen three kinds of silicon solar cells. All of them have different characteristics. The three kinds of silicon cells manufactured during this time were as following:

1. Single crystal silicon cell
2. Multi crystalline silicon cell
3. Thick silicon film

A.DSSC: - This name DSSC stands for "dye sensitized solar cells". A dye-sensitized solar cell (DSSC) is a solar cell that belongs to the group of thin film solar cells. A thin film solar cell works on the principle of photosynthesis in plants. This cell was developed by the scientist known as Michael Gratzel and Brian O Regan at the Ecole Polytechnique Federale de Lausanne in 1991. That's why the cell is known as GRATZEL'S solar cell also. He used mostly artificial dyes. But some organic dyes were also tried by him. The basic idea behind its development was, that every day we can see the most fascinating "solar cells", for example green spinach, algae, and green leaves on trees, converting sunlight by means of photosynthesis into energy containing nutrients like sugar, which are important for life. Why couldn't we use dyes either organic or synthetic dyes with a little bit of sunlight to generate electricity.

II. Manufacturing Steps Of DSSC

A. Negative Electrode: For the making of negative electrode we need to identify first the conducting side of the glass. This has been made conductive by the help of coating layer of Transparent Conductive Oxide (TCO). The coating is done only on one side of the glass so we use multimeter to find out the conducting side of the glass. Then we have to put a thin layer of titanium dioxide on the conducting side. For this we use pipette to put few drops of TiO₂ on the glass surface. And then we use microscope slider glass to spread it through the whole glass surface. It need to be done very carefully then only the result will be good. Otherwise we would not get the required output. After this we dry the layer by the help of Hair Dryer operating in cool mode or we can dry it naturally in room condition.

Fig.1: Putting TiO₂ layer over conductive side of glass

B Dye Solution: The dye solution absorbs most of the sunlight falling on it. This semiconductor TiO₂ layer absorbs only a small fraction of sunlight falling on it (mainly in the UV range). We make solution by the help of different dried fruits or leaves in water. For this we take any of the dye providers and dry them naturally then they become ready to use. We had "Butea frondosa (Palash)" leaves available with us. And we have tried the above given procedure to make dye with them. After putting them in the boiling water, we will observe that they will leave their color with leaving some chlorophyll in the water. Color of water will become same like the fruit or leaf. Then this dye solution will be ready to be use. But for some other natural fruits the procedure of dye making becomes different from the above given.

Fig.2: TiO₂ coated glass in dye solution and Dried Palash Leaves

C Positive Electrodes: To prepare this we need to follow the same procedure first which have been followed in case of negative electrodes. First of all we need to determine the conducting side of the plate by the help of multimeter. After this we put a layer of graphite on the conducting side of the glass. For this we start scratching the Pencil against the conducting side of the electrode. Then a uniform layer of graphite is formed on the surface of conducting glass which works as a catalyst. And it provides the path for the electrons to flow.



Fig.3: Graphite coating over glass plate

The study about the counter electrode has shown if we use a composite plate made with graphite then that the bulk electrical resistance of the composite plate gradually decreases from 6.7 mΩ cm to 1.7 mΩ cm as the graphite content increases in the counter electrode. It happens due to the formation of efficient electronic conducting networks. For DSSCs, the composite plates may be suitable substitutes for the conductive glass plates in the counter electrodes substrates of DSSCs.

Experimental results revealed that composite plates at the optimum level (80 wt.% graphite loading) provide lower cell resistance, lower preparation cost and higher cell performance than common conductive glass plates. Therefore, in order to decrease the cost of the cells and to maintain good cell performance, this graphite-like composite plate prepared is a promising substitute for DSSCs.

D Electrolyte: Here we use Iodide type of electrolyte. This Electrolyte solution is kept in between positive electrode and TiO₂ layer. Electrolyte performs a very important function in the whole electron flow circuit. After the dye molecule has lost one electron the dye molecule will decompose if a electron is not provided quickly. So dye molecule strip the electron by oxidizing iodide electrolyte into tri-iodide. This reaction takes very less time compare to what the lost electron from dye molecule takes to regain its original position. This reduces the chances of short circuit of solar cell. The above given steps make the Dye molecule again capable of providing electron for flowing in the external circuit.

Then this tri-iodide gets its electron by diffusing to the lower area of electrolyte, towards the positive electrode. There the lost electrons come after flowing through external circuit. These electrons combine with the iodide electrolyte. So electrolyte again regains its lost electron. So to close the circuit of electron flow the electrons flow through the negative electrode to the external load then they flow through positive electrode then by help of graphite they come to the electrolyte. This moves them to the negative electrode where dye molecules available in the porous space of TiO₂ reabsorb them. In this manner they again become able to provide electrons for flowing in the external circuit.



Fig 4: Electrolyte Solution

III. Experimental Kit

We had the kit available, which contained all the articles to make a dye sensitized solar cell. It contained following articles.



Fig 5: Experimental kit components

- 1. Glass Plates:** - All the glass plates have transparent conductive oxide (TCO) layer on one side, which make the glasses conductive. Transparent conductive oxide is made of SnO₂ or SnO₂:F.
- 2. TiO₂ solution:** - This solution is given in a certain quantity, to apply it over conductive side of glasses. The solution of TiO₂ powder was made with the acetic acid.

3. Dried Hibiscus leaves are given, which are used as a sensitizer. We need to add these leaves in boiling water to make dye.

4. Electrolyte: - As an electrolyte solution Iodine solution is there. This converts in to the tri-iodide after reacting with graphite which works as a catalyst.

5. Multimeter and wire mesh: - A multimeter is there which is used to measure short circuit current and open circuit voltage, if we connect it directly to the DSSC. It is also used to find out conductive side of the glasses.

6. Pipette, Microscope slider glass, crocodile clips: - We need pipette at the time of coating TiO₂ layer on conductive side.

7. Pencil and safety goggle: - Pencil is used to put a graphite layer on the positive electrode of DSSC and Safety goggle is given to perform furnace related operations safely.

IV. Working Principle

This transforms the energy present in light into electricity by means of following steps: DSSC kit absorbs the sunlight. Then the energy of sunlight is absorbed by the semiconductor material and then this energy is transferred to the dye molecules. It makes dye electrons to move to a higher energy level. Or we can say that electrons move from the valence band to the conduction band. Then it becomes mobile and moves through the titanium dioxide layer. Then it is collected on one side of conducting glass.

After electrons are collected on one side of conducting glass, we connect an external circuit between two electrodes. Electrons start moving through this external circuit connected with external load. Then electrons after moving through external circuit comes back to previous position after moving through counter electrode. This forms a close circuit helping lost electrons to re enter in the dye molecules. This is required for the continuous flow of electricity. Positive electrode collects these electrons from the external circuit and transfers those electrons to ions present in electrolyte. For this we need a catalyst for which graphite is used. This is deposited on the conducting TCO layer of counter electrode.

This catalyst provides help for the flow of electrons. Then the electrons through catalyst reach the electrolyte solution situated in between the positive and negative electrodes. And in the electrolyte solution there are a large number of charged ions. These ions get decomposed and provide electrons to the dye molecule whenever it is needed. The charged ions carry electrons to their original position through the porous areas of titanium dioxide network to the dye molecules where they regain their original positions. This step closes the circuit and recharged dye is capable of repeating process of transforming light into electricity. In this way electrical circuit is completed. Following figure completely depicts the working of these cells.

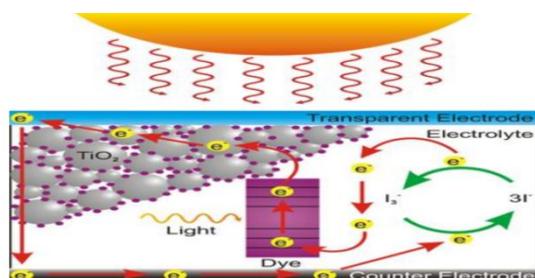


Figure 6: DSSC working principal

V. Experimental Work

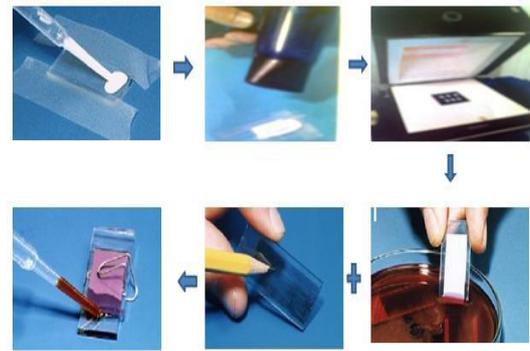


Fig 7: Line diagram for DSSC assembling

1. Putting and spreading TiO₂ solution over conductive glass surface.
2. Drying this layer by help of hair dryer or naturally.
3. Sintering the dried semiconductor layer in the furnace.
4. Putting sintered semiconductor glass plate into dye solution.
5. Preparing positive electrode by adding graphite surface over conductive side of another glass.
6. Keeping both glasses together with the help of clips. Then we add electrolyte solution in- between positive and negative electrodes. After this we leave the cells for 15 to 20 minutes. In this time electrolyte solution spreads through whole semiconductor layer.

A. Formulas Used.

1. $DSSC \text{ EFFICIENCY} = (FF * I_{sc} * V_{oc}) / (\text{SOLAR INPUT})$
Where $FF = \text{Fill Factor}$, $I_{sc} = \text{Short circuit current}$, $V_{oc} = \text{Open circuit voltage}$
2. $FILL \text{ FACTOR} = (I_{max} * V_{max}) / (I_{sc} * V_{oc})$
3. $DSSC \text{ EFFICIENCY} = (I_{max} * V_{max}) / \text{SOLAR INPUT}$

B. Dyes Tried

We have tried various alternatives to make dye sensitizer. For this we have made dyes from following natural fruits and leaves etc.

DYES: -

1. DRIED PALASH (Butea frondosa) FLOWERS
2. DRIED HIBISCUS FLOWERS
3. POMEGRANATE
4. RED WINE
5. BEET
6. MEHANDI
7. TURMERIC etc.

VI. RESULT & DISCUSSION

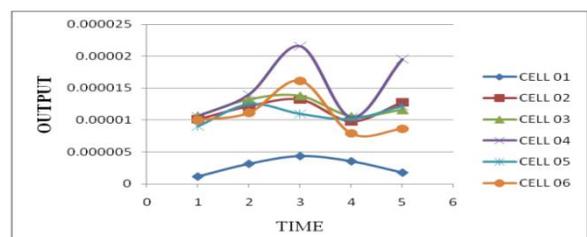


Fig.8: Variation in output of DSSCs with day time for Palash Flowers

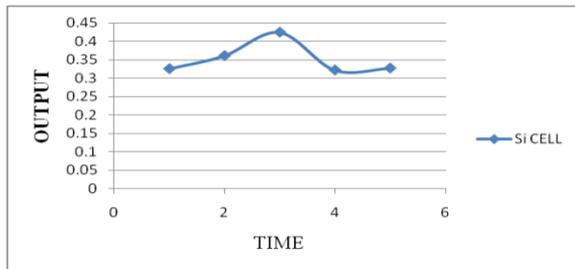


Fig.9: Variation in output of Silicon cell with day time

For the DSSC using dye solution made of Palash flowers, cell 4 has been most efficient in between all DSSCs. The thickness of cell 4 was 0.025 mm. For this dye all the cells have shown a common variation pattern in power output. Power output from the cells first increases with time, as the solar radiation input increases. Then it starts decreasing. And for the silicon cell this pattern is same as that of DSSCs.

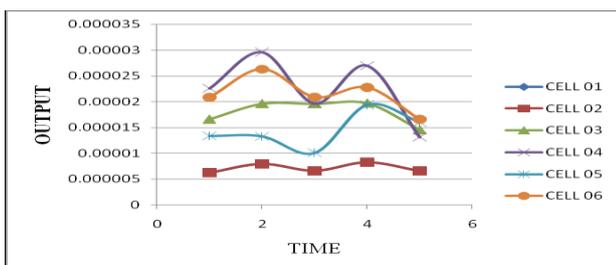


Fig.10: Variation in output of DSSCs with day time for Hibiscus Leaves

For the Hibiscus dye again cell 4, has shown maximum efficiency at each point of time. Here input increases from point 1 to 2, then decreases till point 5. Silicon cell has followed the same pattern as in case of Palash flowers, but DSSCs have shown some irregular pattern. For them output should start decreasing from point 3 to 5. But output did increase from point 3 to 4 then it decreased.

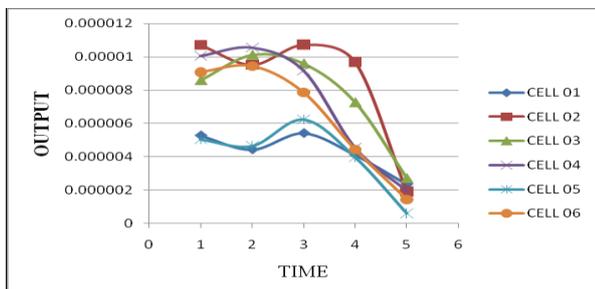


Fig. 11: Variation in output of DSSCs with day time for Pomegranate juice

For the Pomegranate juice all the DSSCs and Silicon cell have indicated the same pattern of output variation with input variation. And cell 2 was the best cell, which had thickness of 0.024mm.

VII. CONCLUSION

By observing all the work that we have done, we can draw following conclusions, regarding dye solutions:

1. DSSCs can work effectively with the natural dyes.
2. Efficiency of DSSC varies with the dye solutions used.
3. All the dye solutions have given different results with compare to other dyes.

4. Out of all eight dyes tried, dye made of DRIED HIBISCUS LEAVES, have given best result.
5. Output variation for all DSSCs with different dyes is showing fluctuating results with respect to time. But in general we can say that Power output of DSSC decreases with time. This is valid for Silicon cell also. With the increase of time the input light intensity first increases up to 12AM, after this intensity starts decreasing. Output from DSSC also first increases then decreases after 12.

VII. FUTURE WORK

We have drawn some conclusion from the project work. But still many areas are there, where we can improve DSSC efficiency.

1. In spite of trying manual coating of TiO₂ layer, other methods like screen printing should be tried. Only by the help of machines, we can get a uniform layer of TiO₂ throughout the area of semiconductor over conductive side of glass. This will help to improve the efficiency of DSSC.
2. At the time of experimentation the input light intensity should be constant. So artificial light source, should also be tried out. This will make us enable to accurately measure the variation of performance of DSSC with the variation of light input.
3. Some other seasonal fruits like Strawberries, Blackberries etc should be tried out.
4. At least one artificial dye should be used to compare their performance with natural dyes.
5. Comparison between all DSSC's performance and their variation will give us effective results, only if all the input parameters for cells are same. Also the construction of all DSSC should be similar to each other. Even all the cells should be facing towards direct sunlight.

REFERENCES

- [1] John T. Sheridan, "Characterising dye-sensitised solar cells". Journal of optic optics, 2010.
- [2] Michael Gratzel, "Review of Dye-sensitized solar cells", (2003) 145–153. Journal of Photochemistry and Photobiology C: Photochemistry Reviews 4 (2003) 145–153.
- [3] Tianyou Peng, Liang Ma and Ke Fan, "Effects of paste components on the properties of screen-printed porous TiO₂ film for dye-sensitized solar cells", Renewable Energy 35 (2010) 555–561.
- [4] Li Yang, "Influence of the preparation conditions of TiO₂ electrodes on the performance of solid-state dye-sensitized solar cells with CuI as a hole collector", Solar Energy 81 (2007) 717–722.
- [5] Jihuai Wu, Yunfang Huang, Sancun Hao, Yelin Wei, "The influence of acid treatment of TiO₂ porous film electrode on photoelectric performance of dyesensitized solar cell", Solar Energy 76 (2004) 745–750.
- [6] M. Giannouli, P. Yianoulis, G. Syrokostas "Effects of paste storage on the properties of nanostructured thin films for the development of dye-sensitized solar cells", Renewable Energy 34 (2009) 1759–1764.
- [7] Jihuai Wu, Jianming Lin, Sancun Hao, Yunfang Huang "Natural dyes as photosensitizers for dye-sensitized solar cell", Solar Energy 80 (2006) 209–214.
- [8] Eiji Yamazaki, Masaki Murayama, Naomi Nishikawa, Noritsugu Hashimoto, Masashi Shoyama, "Utilization of natural carotenoids as photosensitizers for dyesensitized solar cells", Solar Energy 81 (2007) 512–516.
- [9] Kuo-Chuan Ho, Chun-Guey Wu, Yung-Liang Tung, Shi-Jhang Wu, Chia-Yuan Chen, "An efficient light-harvesting ruthenium

dye for solar cell application”, Dyes and Pigments 84 (2010) 95–101.

- [10] Yong Soo Kang, Jong Hak Kim, Rajkumar Patel, Jin Ah Seo, Joo Hwan Koh. “Dye-sensitized solar cells employing amphiphilic poly(ethylene glycol) electrolytes”, Journal of Photochemistry and Photobiology A: Chemistry 217 (2011) 169–176.

AUTHORS PROFILE



Mr. Thirumal Mamidi received his M.Tech (Heat Power Engineering) degree from Visveswaraya National Institute Of Technology (VNIT) Nagpur In 2011.He is working as Asst.Professor in the M.B.E.S.College of Engineering Ambajogai-431517,Beed,Maharashtra.

Abhishek Mishra received his M.Tech (Heat Power Engineering) degree from Visveswaraya National Institute Of Technology (VNIT)Nagpur In 2011.He is working in Tata Engineering Services (TES) Jharkhand.

Transmission Loss Minimization using UPFC

Banakar Basavaraj¹, Ronad Basangouda², Jangamshetti Suresh.H³

¹Student Member IEEE, ²Member IEEE, ³Senior Member IEEE

Abstract: The paper focuses on the issue of transmission loss allocation and transmission loss minimization. To meet the power demand it is essential to increase the transmitted power by reducing transmission losses. This can be achieved by adding Flexible AC Transmission System (FACTS) controllers. The annual, weighted average transmission and distribution loss in India is about 30-35%. To investigate the effect of the UPFC on the steady state condition of the system, different models can be used. The Unified Power Flow Controller (UPFC) injection model is incorporated in load flow by Newton Raphson algorithm. Loss allocation is an important aspect in determining the cost of transmission. Z-bus loss allocation technique is used to achieve the same. Currents rather than powers are emphasized in the allocation process. Optimal location to place UPFC is identified based on active power loss Sensitivity factors. The changes in the system are studied to see the impact of the UPFC. The impact of UPFC is analyzed by using 5-bus test system, IEEE 14-bus and 30-bus test systems. The results conclude that considerable amount of losses can be reduced. The analysis is achieved through developing a program using MATLAB.

Index Terms: LFA (load flow analysis), UPFC (unified power flow control), Z bus Allocation, Sensitivity factors, Flexible AC transmission systems (FACTS).

I. INTRODUCTION

In the present pace of power system, transmission systems are being required to provide increased power transfer capability to accommodate a much wider range of possible generation patterns. Environmental, right-of-way, and cost problems are major hurdles for power transmission network expansion. Hence, there is an interest in better utilization of available power system capacities by installing FACTS controllers. Power systems today are highly complex in nature. The requirements to provide a stable, secure, controlled quality of power are becoming vitally important with the rapid growth in industrial area. To meet the demand in a power system it is essential to increase the transmitted power by reducing losses in the transmission line. Moreover installation of new transmission lines in a power system leads to the technological complexities increased cost and delay in construction. Considering these factors it is necessary to modify the existing transmission system instead of constructing new transmission lines.

II. UNIFIED POWER FLOW CONTROLLER (UPFC)

A. Circuit Arrangement of UPFC

The UPFC consists of two voltage source converters, which are connected back to back through a DC link. The series voltage converter is connected to the transmission

line by means of a series transformer and the shunt voltage converter by means of shunt transformer. The series voltage converter injects an AC voltage into the transmission line with Controllable magnitude and phase angle.

The shunt converter can exchange active and reactive powers with the system, which enables the system to do shunt compensation independently. Converter 2 provides the main function of the UPFC by injecting an AC voltage with controllable magnitude and phase angle in series with the transmission line via a series transformer. The basic function of converter 1 is to supply or absorb the real power demand by converter 2 at the common DC link. It can also generate or absorb controllable reactive power and provide independent shunt reactive compensation for the line. Converter 2 supplies or absorbs locally the required reactive power and exchanges the active power as a result of the series injection voltage [2].

Converter 2 is used to generate a voltage source at the fundamental frequency with variable amplitude and phase angle, which is added to the ac transmission line by the series-connected boosting transformer. The inverter output voltage injected in series with line can be used for direct voltage control, series compensation, phase shifter, and their combinations.

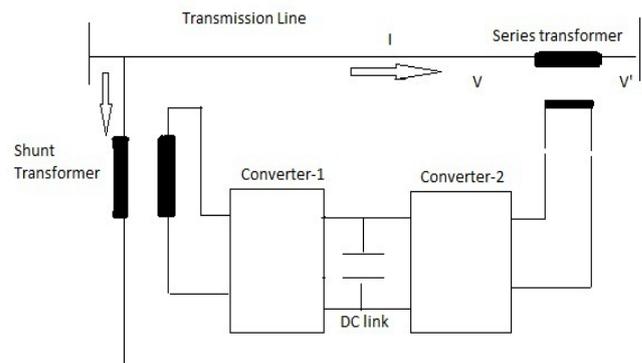


Fig.1. Schematic diagram of UPFC [1].

Converter 2 is used to generate a voltage source at the fundamental frequency with variable amplitude and phase angle, which is added to the ac transmission line by the series-connected boosting transformer. The inverter output voltage injected in series with line can be used for direct voltage control, series compensation, phase shifter and their combinations. This voltage source can internally generate or absorb all the reactive power required by the different type of controls applied and transfers active power at its dc terminal.

B. UPFC injection model for power flow studies.

Model for UPFC, which will be referred as UPFC injection model is derived. This model is helpful in understanding the impact of the UPFC on the power system in the steady state. Further, the UPFC injection model can easily be

incorporated in the steady state power flow model. The series voltage source converter does the main function of the UPFC by injecting voltage in series with transmission line [1]. Series connected voltage source converter model: Suppose a series connected voltage source is located between nodes i and j in a power system. The series voltage source converter can be modeled with an ideal series voltage V_s in series with a reactance X_s as shown in fig below.

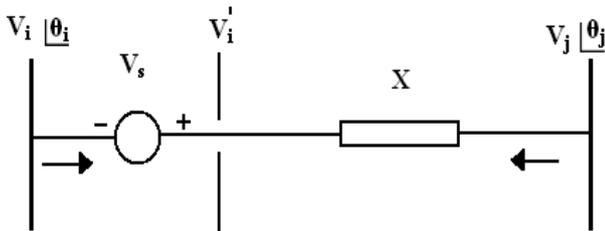


Fig.2. Representation of a series connected VSC

$$V_i^1 = V_s + V_i \quad (1)$$

where

V_i^1 = voltage behind the series reactance.

V_s = series source voltage.

V_i = voltage at i'th node.

The series voltage source V_s is controllable in magnitude and phase

$$V_s = rV_i e^{j\gamma} \quad (2)$$

where r = series voltage source co-efficient. ($0 < r < r_{max}$)

γ = series voltage source angle. ($0 < \gamma < 2\pi$)

The injection model is obtained by replacing the equivalent circuit of series connected voltage source as Norton's equivalent circuit as shown in fig.5 .The current source,

$$I_s = -jb_s V_s \quad (3)$$

Where,

$$b_s = 1 / X_s$$

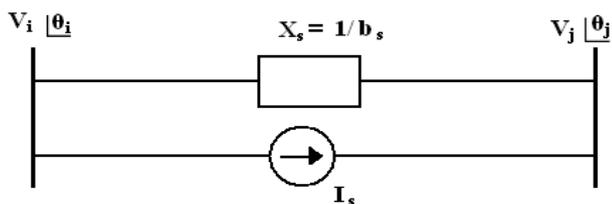


Fig.3. Equivalent Norton's circuit of a series connected VSC.

The power injected into the jth bus

$$\overline{S_{js}} = \overline{V_j} (-\overline{I_s})^* \quad (5)$$

$$S_{js} = V_j [-jb_s r \overline{V_i} e^{j\gamma}]^*$$

$$S_{js} = b_s r V_i V_j \sin(\theta_{ij} + \gamma) + jb_s r V_i V_j \cos(\theta_{ij} + \gamma)$$

Where $\theta_{ij} = \theta_i - \theta_j$

From above equations, the injection model of series connected voltage source can be sent as two dependent loads as shown in fig.4. In UPFC, the shunt connected voltage source (converter1) is used mainly to provide the active power, which is injected to the network via the series connected voltage source.

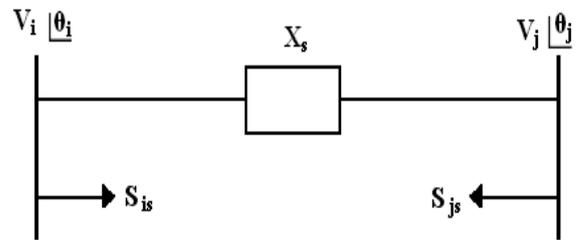


Fig.4. Injection model for a series connected VSC.

When the losses are neglected [1]

$$P_{conv1} = P_{conv2}$$

The apparent power supplied by the series voltage source converter is

$$S_{conv2} = \overline{V_s} \overline{I_{ij}}^* = r e^{j\gamma} \overline{V_i} \left[\frac{\overline{V_i^1} - \overline{V_j}}{jX_s} \right]^* \quad (6)$$

After simplification, the active and reactive power supplied by converter 2 is

$$P_{conv2} = rb_s V_i V_j \sin(\theta_i - \theta_j + \gamma) - rb_s V_i^2 \sin(\gamma)$$

$$Q_{conv2} = -rb_s V_i V_j \cos(\theta_i - \theta_j + \gamma) + rb_s V_i^2 \cos(\gamma) + r^2 b_s V_i^2 \quad (7)$$

The reactive power delivered or absorbed by converter 1 is independently controllable by UPFC and can be modelled as a separate controllable shunt reactive source. In view of above, it is assumed that $Q_{conv1} = 0$. The UPFC injection model is constructed from the series connected voltage source model with the addition of a power equivalent to $P_{conv1} + j0$ to node i.

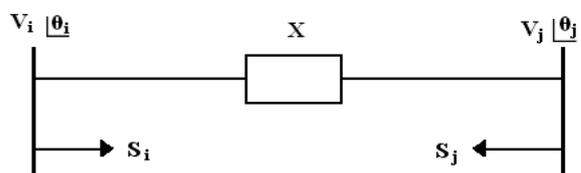


Fig.7. Complete UPFC model.

Thus, the complete UPFC injection model is shown in fig.7.

$$\begin{aligned}
 Q_{si} &= rb_s V_i^2 \cos(\gamma) + Q_{shunt} \\
 Q_{sj} &= -rb_s V_i V_j \cos(\theta_{ij} + \gamma) \\
 P_{si} &= rb_s V_i V_j \sin(\theta_{ij} + \gamma) \\
 P_{sj} &= -rb_s V_i V_j \sin(\theta_{ij} + \gamma)
 \end{aligned}
 \tag{8}$$

III. Methodology

A. Incorporation of UPFC in Newton Raphson Power flow algorithm:

The Steps to Incorporate UPFC in Newton-Raphson Algorithm are as follows

- Read the system input data; line data, bus data, generator and load data.
- Formation of admittance matrix 'Y' bus of the transmission line between the bus i and j.
- Combining the UPFC power equations with network equation, the conventional power flow equation is given as:

$$P_i + jQ_i = \sum_{j=1}^n V_i V_j Y_{ij} \angle(\theta_{ij} - \delta_i + \delta_j) + P_i^1 + jQ_i^1
 \tag{9}$$

- The conventional jacobian matrix are formed. Due to the inclusion of UPFC the dimensions of the jacobian matrix is increased.
- In this step, the jacobian matrix is modified and power equations are mismatched. The Bus bar voltages are updated at each iteration and convergence is checked. If convergence is not achieved in the next step the algorithm goes back to the step 5 and the jacobian matrix is modified and the power equations are mismatched until convergence is attained.
- If the convergence achieved in earlier step, the output load flow is calculated for PQ bus that includes the Bus bar voltages, generation, transmission line flow and losses.

B. Sensitivity analysis of total active power loss

A method based on the sensitivity of the total system active power loss with respect to the control variables of the FACTS device UPFC is considered. For UPFC placed between buses i and bus j, the considered control parameter is the injected series voltage of controllable magnitude and its phase angle. The active power loss sensitivity factor with respect to these control variables may be given as, loss sensitivity with respect to control parameter of UPFC placed between buses i and bus j[3].

$$a_{ij} = \frac{\partial P_L}{\partial V_{ij}}
 \tag{10}$$

And this can be deduced from the above equation as,

$$\frac{\partial P_L}{\partial V_{ij}} = 2V_i V_j \cos(\delta_i - \delta_j) + 2V_i V_j \sin(\delta_i - \delta_j)
 \tag{11}$$

Thus from the above equation the sensitivity factors with respect to active power loss are obtained.

C. Z-Bus loss allocation method

The goal of the Z-bus loss allocation method, is to take a solved power flow and systematically distribute the system transmission losses, among the network buses. The loss component, L_k is the fraction of the system losses allocated to the net real power injection at bus K[4]. The Z-bus loss allocation algorithm is as follows

- Solve load flow; get bus voltage vector V and total power loss.
- Obtain bus current vector I from V and complex power injection.

$$S = (P_i + jQ_i)
 \tag{12}$$

- Obtain the vector RI.

$$RI = \text{Re} \{Z\} I.$$

$$RI = \text{Re} \{Z [\text{Re} (I)]\} + \text{Re} \{Z [\text{Im} (I)]\}
 \tag{13}$$

- Calculate the component of total loss due to current injection I_k at the jth bus.

$$L_k = \Re \left[I_k^* \left(\sum_{j=1}^{nb} R_{kj} \cdot I_j \right) \right]
 \tag{14}$$

IV. SIMULATION RESULTS

The simulation is carried out for IEEE 14 and 30 bus systems.

A. Simulation result for IEEE-14 bus system.

Table .1.Voltage profile & distribution of active power Loss cost without UPFC.

Bus Number	Voltage (p.u.)	Angle (degree)	Distribution of active power loss in \$/hr	Distribution of active power loss in ₹/hr
1	1.0600	0.000	388	19400
2	1.0350	-4.8589	7	350
3	0.9922	-12.6963	144	7200
4	1.0026	-10.2365	45	2250
5	1.0082	-8.7363	4	200
6	1.0610	-14.3841	6	300
7	1.0313	-13.3617	0	0
8	1.0313	-13.3711	0	0
9	1.0357	-14.9825	27	1350
10	1.0371	-15.2613	9	450
11	1.0471	-14.9944	3	150
12	1.0511	-15.3826	6	300
13	1.0471	-15.5190	15	750
14	1.0299	-16.4270	22	1100
			677	33,850

Table 1 depicts the voltage, phase angle and distribution of active power loss cost in the system before incorporating UPFC in the system. Before incorporating UPFC in the system, the total real power loss in the system is 13.5272 MW. The generator 1 (i.e., at Bus-1) with about 85% of the total generation always gets allocated with highest cost,

which is proportional to the highest loss taking place in that Bus-1(slack bus generator).

Since, the sensitivity factors are calculated with respect to the active power loss in the system the sensitivity factor at that Bus-1 will be reasonably high as it can be seen in the Table 2. The sensitivity factors from the Table 2 indicate that the line-4, line-5, and line-10 are selected as the best lines for locating UPFC based on their highest sensitivity factors.

Table.2. Sensitivity factors:

Bus Number	UPFC in line-4		UPFC in line-5	
	Voltage (p.u)	Angle (degree)	Voltage (p.u)	Angle (degree)
1	1.0600	0.0000	1.0600	0.0000
2	1.0350	-4.6739	1.0350	-4.8228
3	0.9922	-12.5313	0.9923	-12.6407
4	1.0026	-10.0904	1.0028	-10.1648
5	1.0083	-8.6033	1.0085	-8.6533
6	1.0610	-14.2468	1.0612	-14.3026
7	1.0313	-13.2177	1.0315	-13.2864
8	1.0313	-13.2272	1.0315	-13.2959
9	1.0357	-14.8397	1.0359	-14.9055
10	1.0372	-15.1194	1.0373	-15.1833
11	1.0471	-14.8548	1.0473	-14.9146
12	1.0511	-15.2449	1.0513	-15.3010
13	1.0471	-15.3809	1.0473	-15.4377
14	1.0299	-16.2863	1.0301	-16.3476

Table .3. Comparisons of voltage and phase angle profile with UPFC.

Bus Number	UPFC in Line-4	UPFC in Line-5
	Distribution of active power loss cost in \$ / hr	Distribution of active power loss cost in \$ / hr
1	369.1634	382.2617
2	10.1278	7.2548
3	143.4611	144.1806
4	44.4162	44.1990
5	3.9037	3.0096
6	6.0450	5.9372
7	0.0000	-0.0000
8	0.0919	0.0914
9	26.9393	26.7652
10	8.9313	8.8719
11	2.9764	2.9481
12	5.7665	5.7077
13	14.5446	14.4183
14	22.4375	22.3233
Total cost	659	668
Total loss	13.1968 MW	13.3774 MW

Table 3 shows the comparisons of voltage and phase angle profiles after incorporating UPFC in the line-4, line-5 and line-10. From the Table 3 it can be seen that the voltage and phase angles has been improved considerably better in the selected line-4, line-5 and line-10 than compared to without UPFC results as shown in Table 1

The Table 4 shows the comparisons of loss allocation costs at each bus for the three lines namely line-4, line-5 and line-10. On Comparison, it is found that the system losses have considerably reduced. Thus, placing UPFC in line-4 gives efficient and better results than in other lines. The reduction in losses is 0.3304 MW/hr its respective cost reduction is 17 \$/hr [3].

Table.4. Comparisons of Loss minimization and Loss allocation with UPFC

Line Number	From Bus	To Bus	Sensitivity factor
1	2	4	2.2607
2	2	5	2.2234
3	2	3	2.3147
4	1	2	2.3722
5	1	5	2.4373
6	3	4	1.9023
7	5	4	2.0740
8	4	7	2.1776
9	4	9	2.2415
10	5	6	2.3397
11	6	11	2.2454
12	6	12	2.2689
13	6	13	2.2654
14	7	8	2.1275
15	7	9	2.1958
16	9	10	2.1588
17	9	14	2.1865
18	10	11	2.1618
19	12	13	2.2063
20	13	14	2.1906

V. CONCLUSION

REWRITE THE CONCLUSION. Include the results of the 30 bus system in conclusion. Once again check the technical part of the paper thoroughly. The results are obtained for IEEE 14-bus and 30-bus test system. The transmission losses are reduced satisfactorily. UPFC is optimally placed using active power loss sensitivity factors which were calculated after performing the load flow analysis. UPFC's role in loss minimization and its influence for loss allocation is verified. Minimization of power losses can be achieved with out generation scheduling. By controlling the angle and magnitude of injected voltage, power flow through transmission line can be controlled.

REFERENCES

[1] Loss minimization by incorporation of UPFC in load flow studies, s.v. ravi kumar and s. Siva nagaraju. International journal of electrical and power engineering vol1(3):321-327.
 [2] Steady State Analysis of Unified Power Flow Controller: Mathematical Modeling and Simulation Studies A.M. Vural, Student Member, IEEE, and M. Tumay.
 [3] Optimal location of unified power flow controller (UPFC) in Nigerian Grid System, Mark Ndubuka Nwohu, International journal of electrical and Power Engineering, medwell journals,2010.

- [4] A.J Conejo et al, "Z Bus Loss Allocation," IEEE Trans.Power Syst., vol 16 no.1, pp.105-110, feb,2001.
- [5] Study and Effects of UPFC and its Control System for Power Flow Control and Voltage Injection in a Power System, Vibhor Gupta / International Journal of Engineering Science and Technology Vol. 2(7), 2010, 2558-2566.
- [6] Analysis and Modeling of UNIFIED POWER FLOW CONTROLLER: Modification of NEWTON-RAPHSON Algorithm and user-defined modeling approach for POWER FLOW STUDIES, M. Tumay and A. M. Vural, October 2004 The Arabian Journal for Science and Engineering, Volume 29, Number 2B.
- [7] L.Gyugyi, "A unified power flow control concept for flexible AC transmission systems," Proc. Inst.Elect.Eng. vol. 39, no. 4, pp. 323-331, Jul. 1992.
- [8] A unified power flow controller concept for flexible AC transmission systems," in Proc. 5th Int. Conf. AC DC Transmission, London, U.K., Sep. 17-20, 1991 vol. 19-26.

BIOGRAPHY

Basavaraj Bankar. was born in Karnataka, India on 17 February 1988. HE obtained B.E (Electrical) from Visveshwaraiah technological University, Karnataka.,India. He is currently persuing M.Tech Degree in Power and Energy Systems in Electrical and Electronics Engineering, Basaveshwar Engineering College, Bagalkot, India.

Basangoud.F.Ronad Was born in Badami, Karnataka, India on 22 July 1986. He obtained B.E (Electrical and Electronics) from Visveswaraih Technological University, Belgaum, Karnataka India in 2008 and M.Tech (Power and Energy System) from Basaveshwar Engineering College, Bagalkot, Karnataka, India in 2010. His areas of interest include Power Electronics, Power system and Renewable Energy Sources. Presently he is working as faculty in the Department of Electrical & Electronics Engineering at Basaveshwar Engineering College, Bagalkot, India.

Dr. Suresh. H. Jangamshetti: (S'88, M'90, SM'97) was born in Bijapur, Karnataka, India on May 28, 1963. He obtained his B.E (Electrical) degree from Karnataka University Dharwad in 1985 and M.Tech. (Power Systems) & Ph.D (Wind Energy Systems) from IIT Kharagpur in 1989 & 2000 respectively.

His areas of interest include Wind-Solar Energy Systems, Energy Conservation, Computer Applications to Power System and FACTS He won the "Outstanding IEEE Student Branch Counsellor" award for the year 1996(R10) and 2010 (IEEE Bangalore Section) at Basaveshwar Engineering College, Bagalkot, Karnataka, India. He was Fulbright-Nehru Visiting Lecture Fellow at Michigan Technological University, Houghton MI USA during Fall 2011. He is working as Professor in the department of E&E at Basaveshwar Engineering College, Bagalkot.

A new form of weaker separation axioms via $pgr\alpha$ -closed sets

K. Binoy Balan¹ and C. Janaki²

¹Asst. Professor, Dept. of Mathematics, Sree Narayana Guru College, Coimbatore- 105, India.

²Asst. Professor, Dept. of Mathematics, L.R.G. Govt. Arts College for Women, Tirupur-4, India.

Abstract: The aim of this paper is to introduce and characterize $pgr\alpha$ - regular spaces and $pgr\alpha$ - normal spaces via the concept of $pgr\alpha$ - closed sets. It also focuses on some of its basic properties and discusses on separation axioms between $pgr\alpha$ - T_0 and $pgr\alpha$ - T_1 . An attempt has been made to make a comparative study with other usual separation axioms.
 Mathematics Subject Classification: 54C10, 54C08, 54C05

Key Words: $pgr\alpha$ - T_0 , $pgr\alpha$ - T_1 , $pgr\alpha$ - T_2 , $pgr\alpha$ - $T_{1/2}$, $pgr\alpha$ - $T_{1/3}$, $pgr\alpha$ - T_b , $pgr\alpha$ - $T_{3/4}$, $pgr\alpha$ - normal and $pgr\alpha$ - regular spaces.

I. Introduction

Separation axioms in topological spaces play a dominant role in analysis and are usually denoted with the letter “T” after the German “Training” which mean separation. The separation axioms that were studied together in this way were the axioms for Hausdorff spaces, regular spaces and normal spaces. Separation Axioms and closed sets in topological spaces have been very useful in the study of certain objects in digital topology [3,5]. Khalimsky, Kopperman and Meyer [4] proved that the digital line is a typical example of $T_{1/2}$ spaces. The first step of generalized closed set was done by Levine [6] in general topology which was properly placed between T_0 - space and T_1 -space. After the works of Levine on semi open sets, several mathematicians turned their attention to the generalization of various concepts of topology. Consequently, many separation axioms has been defined and studied. We introduce a weaker form of separation axioms called $pgr\alpha$ - separation axioms using the concept of $pgr\alpha$ - open sets introduced in [1]. In this paper the concepts of $pgr\alpha$ - T_0 , $pgr\alpha$ - T_1 , $pgr\alpha$ - regular and $pgr\alpha$ -normal are introduced and basic properties are discussed.

II. Preliminaries

Throughout this paper (X, τ) represents nonempty topological spaces on which no separation axioms is assumed unless otherwise mentioned. For a subset A of a topological space X, $cl(A)$ and $int(A)$ denote the closure of A and the interior of A respectively. In this section, some definitions and theorems are further investigated which are used in this work.

Definition 2.1 A subset A of a space (X, τ) is called

- (i) a pre open set [7] if $A \subset int(cl(A))$ and a pre closed set if $cl(int(A)) \subset A$.
- (ii) a α -open set [8] if $A \subset int(cl(int(A)))$ and α -closed set if $cl(int(cl(A))) \subset A$.
- (iii) a regular open set if $A = int(cl(A))$ and a regular closed set if $A = cl(int(A))$.
- (iv) a regular α - open set (briefly α - open) [10] if there is a regular open set U such that $U \subset A \subset \alpha cl(U)$.

The union of all pre open sets of X contained in A is called pre- interior of A and is denoted $pint(A)$. Also the intersection of all pre- closed subsets of X containing A is called pre- closure of A and is denoted by $pcl(A)$. Note that $pcl(A) = A \cup cl(int(A))$ and $pint(A) = A \cap int(cl(A))$.

Definition 2.2 A subset A of a space (X, τ) is called

- (i) a generalized closed set (briefly g-closed) [6] if $cl(A) \subset U$ whenever $A \subset U$ and U is open.
- (ii) a generalized α -closed set (briefly $g\alpha$ - closed) [6] if $\alpha cl(A) \subseteq U$ whenever $A \subseteq U$ and U is α - open in X.
- (iii) a generalized pre regular closed set (briefly gpr- closed) [2] if $pcl(A) \subseteq U$ whenever $A \subseteq U$ and U is regular open in X
- (iv) a $pgr\alpha$ - closed set [1] if $pcl(A) \subset U$ whenever $A \subset U$ and U is regular α -open.

The complement of the above mentioned closed sets are their respective open sets.

Definition 2.3 A function $f : X \rightarrow Y$ is called

- (i) $pgr\alpha$ - continuous [1] if for every closed set V of Y then $f^{-1}(V)$ is $pgr\alpha$ - closed set in X.
- (ii) pre- continuous [7] if for every closed set V of Y then $f^{-1}(V)$ is pre-closed set in X.
- (iii) regular- continuous [9] if for every closed set V of Y then $f^{-1}(V)$ is regular closed set in X.
- (iv) gpr-continuous [2] if for every closed set V of Y then $f^{-1}(V)$ is gpr-closed set in X.
- (v) $g\alpha$ - continuous [6] if for every closed set V of Y then $f^{-1}(V)$ is $g\alpha$ - closed set in X.

Definition 2.4 [1] A function $f: X \rightarrow Y$ is $pgr\alpha$ - irresolute if for every $pgr\alpha$ - open set V of Y then $f^{-1}(V)$ is $pgr\alpha$ -open set in X.

Definition 2.5 A space (X, τ) is called a $T_{1/2}$ space [6] ($pgr\alpha$ - $T_{1/2}$ space [1]) if every g-closed (resp. $pgr\alpha$ - closed) is closed (resp. pre closed).

Theorem 2.6 [1] A space X is $pgr\alpha$ - $T_{1/2}$ if and only if every singleton set is regular closed or pre open.

III. On $\text{pgr}\alpha$ - T_k Spaces ($K=0, 1, 2, B, 1/2, 1/3, 3/4$)

Definition 3.1 A topological space X is called

- (i) a $\text{pgr}\alpha$ - T_0 if for each pair of distinct points x, y of X , there exists a $\text{pgr}\alpha$ - open sets G in X containing one of them and not the other.
- (ii) a $\text{pgr}\alpha$ - T_1 if for each pair of distinct points x, y of X there exists two $\text{pgr}\alpha$ - open sets G_1, G_2 in X such that $x \in G_1, y \notin G_1$, and $y \in G_2, x \notin G_2$.
- (iii) a $\text{pgr}\alpha$ - T_2 ($\text{pgr}\alpha$ - Hausdorff) if for each pair of distinct points x, y of X there exists distinct $\text{pgr}\alpha$ - open sets H_1 and H_2 such that H_1 containing x but not y and H_2 containing y but not x .

Theorem 3.2

- (i) Every T_0 -space is $\text{pgr}\alpha$ - T_0 space.
- (ii) Every T_1 - space is $\text{pgr}\alpha$ - T_0 space.
- (iii) Every T_1 space is $\text{pgr}\alpha$ - T_1 space.
- (iv) Every T_2 space is $\text{pgr}\alpha$ - T_2 space.
- (v) Every $\text{pgr}\alpha$ - T_1 space is $\text{pgr}\alpha$ - T_0 space.
- (vi) Every $\text{pgr}\alpha$ - T_2 space is $\text{pgr}\alpha$ - T_1 space.

Proof: Straight forward.

The converse of the theorem need not be true as in the examples.

Example 3.3 Let $X = \{a, b, c\}$ and $\tau = \{X, \phi, \{a\}, \{b\}, \{a, b\}\}$.

Then $\text{PGR}\alpha C(X) = \{\phi, X, \{a, b\}, \{c\}, \{a, c\}, \{b, c\}\}$.

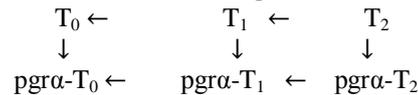
Here (X, τ) is $\text{pgr}\alpha$ - T_0 space but not T_0 space and $\text{pgr}\alpha$ - T_1 space.

Example 3.4 Let $X = \{a, b, c\}$ and $\tau = \{X, \phi, \{a, b\}\}$.

Then $\text{PGR}\alpha C(X) = \{X, \phi, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}\}$. Here (X, τ) is $\text{pgr}\alpha$ - T_1 space but not T_1 space and $\text{pgr}\alpha$ - T_2 space.

Example 3.5 Let $X = \{a, b, c\}$ and τ is indiscrete topology on X , then (X, τ) is $\text{pgr}\alpha$ - T_2 but not T_2 space.

The following diagram shows the relation between usual separation axiom and $\text{pgr}\alpha$ - separation axiom.



Theorem 3.6 Let X be a topological space and Y is an $\text{pgr}\alpha$ - T_2 space. If $f: X \rightarrow Y$ is injective and $\text{pgr}\alpha$ - irresolute then X is $\text{pgr}\alpha$ - T_2 space.

Proof: Suppose $x, y \in X$ such that $x \neq y$. Since f is injective then $f(x) \neq f(y)$.

Since Y is $\text{pgr}\alpha$ - T_2 space then there are two $\text{pgr}\alpha$ - open sets U and V in Y such that $f(x) \in U, f(y) \in V$ and $U \cap V = \phi$.

Since f is $\text{pgr}\alpha$ - irresolute then $f^{-1}(U), f^{-1}(V)$ are two $\text{pgr}\alpha$ - open sets in $X, x \in f^{-1}(U), y \in f^{-1}(V), f^{-1}(U) \cap f^{-1}(V) = \phi$. Hence X is $\text{pgr}\alpha$ - T_2 space.

Theorem 3.7 Let X be a topological space and Y is an T_2 space .If $f: X \rightarrow Y$ is injective and $\text{pgr}\alpha$ - continuous then X is $\text{pgr}\alpha$ - T_2 space.

Proof: Suppose $x, y \in X$ such that $x \neq y$. Since f is injective, then $f(x) \neq f(y)$.

Since Y is an T_2 space, then there are two open sets U and V in Y such that $f(x) \in U, f(y) \in V$ and $U \cap V = \phi$. Since f is $\text{pgr}\alpha$ - continuous then $f^{-1}(U), f^{-1}(V)$ are two $\text{pgr}\alpha$ - open sets in X . Then $x \in f^{-1}(u), y \in f^{-1}(V), f^{-1}(U) \cap f^{-1}(V) = \phi$. Hence X is $\text{pgr}\alpha$ - T_2 space.

Definition 3.8 The intersection (resp. union) of all $\text{pgr}\alpha$ - closed (resp. $\text{pgr}\alpha$ -open) sets each containing in (resp. contained) a set A in a space X is called $\text{pgr}\alpha$ - closure (resp. $\text{pgr}\alpha$ -interior) of A and it is denoted by $\text{pgr}\alpha\text{-cl}(A)$ (resp. $\text{pgr}\alpha\text{-int}(A)$).

Remark 3.9 Let X be a topological space such that $A \subset X$ then $\text{pgr}\alpha\text{-cl}(A)$ is contained in every $\text{pgr}\alpha$ -closed set containing A .

Theorem 3.10 Let X be a topological space and $A \subset B \subset X$ then

- (i) $\text{pgr}\alpha\text{-cl}(A)$ is the smallest $\text{pgr}\alpha$ -closed set which contains A .
- (ii) $\text{pgr}\alpha\text{-cl}(A) \subset \text{pgr}\alpha\text{-cl}(B)$.
- (iii) A is an $\text{pgr}\alpha$ - closed set if and only if $\text{pgr}\alpha\text{-cl}(A) = A$.
- (iv) $\text{pgr}\alpha\text{-cl}(\text{pgr}\alpha\text{-cl}(A)) = \text{pgr}\alpha\text{-cl}(A)$

Theorem 3.11 (X, τ) is $\text{pgr}\alpha$ - T_0 space if and only if for each pair of distinct x, y of $X, \text{pgr}\alpha\text{-cl}(\{x\}) \neq \text{pgr}\alpha\text{-cl}(\{y\})$.

Proof: Let (X, τ) be a $\text{pgr}\alpha$ - T_0 space. Let $x, y \in X$ such that $x \neq y$, then there exists a $\text{pgr}\alpha$ - open set V containing one of the points but not the other, say $x \in V$ and $y \notin V$. Then V^c is a $\text{pgr}\alpha$ - closed containing y but not x . But $\text{pgr}\alpha\text{-cl}(\{y\})$ is the smallest $\text{pgr}\alpha$ - closed set containing y .

Therefore $\text{pgr}\alpha\text{-cl}(\{y\}) \subset V^c$ and hence $x \notin \text{pgr}\alpha\text{-cl}(\{y\})$.

Thus $\text{pgr}\alpha\text{-cl}(\{x\}) \neq \text{pgr}\alpha\text{-cl}(\{y\})$.

Conversely, suppose $x, y \in X, x \neq y$ and $\text{pgr}\alpha\text{-cl}(\{x\}) \neq \text{pgr}\alpha\text{-cl}(\{y\})$. Let $z \in X$ such that $z \in \text{pgr}\alpha\text{-cl}(\{x\})$ but $z \notin \text{pgr}\alpha\text{-cl}(\{y\})$. If $x \in \text{pgr}\alpha\text{-cl}(\{y\})$ then $\text{pgr}\alpha\text{-cl}(\{x\}) \subset \text{pgr}\alpha\text{-cl}(\{y\})$ and hence $z \in \text{pgr}\alpha\text{-cl}(\{y\})$. This is a contradiction. Therefore $x \notin \text{pgr}\alpha\text{-cl}(\{y\})$. That is $x \in (\text{pgr}\alpha\text{-cl}(y))^c$.

Therefore $(\text{pgr}\alpha\text{-cl}(\{y\}))^c$ is a $\text{pgr}\alpha$ - open set containing x but not y . Hence (X, τ) is $\text{pgr}\alpha$ - T_0 space.

Theorem 3.12 A topological space X is $\text{pgr}\alpha$ - T_1 space if and only if for every $x \in X$ singleton $\{x\}$ is $\text{pgr}\alpha$ - closed set in X .

Proof: Let X be $\text{pgr}\alpha$ - T_1 space and let $x \in X$, to prove that $\{x\}$ is $\text{pgr}\alpha$ -closed set. We will prove $X - \{x\}$ is $\text{pgr}\alpha$ -open set in X . Let $y \in X - \{x\}$, implies $x \neq y \in X$ and since X is $\text{pgr}\alpha$ - T_1 space then their exist two $\text{pgr}\alpha$ -open sets G_1, G_2 such that $x \notin G_1, y \in G_2 \subseteq X - \{x\}$.

Since $y \in G_2 \subseteq X - \{x\}$ then $X - \{x\}$ is $\text{pgr}\alpha$ -open set. Hence $\{x\}$ is $\text{pgr}\alpha$ -closed set.

Conversely, Let $x \neq y \in X$ then $\{x\}, \{y\}$ are $\text{pgr}\alpha$ -closed sets. That is $X - \{x\}$ is $\text{pgr}\alpha$ -open set.

Clearly, $x \notin X - \{x\}$ and $y \in X - \{x\}$. Similarly $X - \{y\}$ is $\text{pgr}\alpha$ -open set, $y \notin X - \{y\}$ and $x \in X - \{y\}$. Hence X is $\text{pgr}\alpha$ - T_1 space.

Theorem 3.13 For a topological space (X, τ) , the following are equivalent

(i) (X, τ) is $\text{pgr}\alpha$ - T_2 space.

(ii) If $x \in X$, then for each $y \neq x$, there is a $\text{pgr}\alpha$ -open set U containing x such that $y \notin \text{pgr}\alpha\text{-cl}(U)$

Proof: (i) \Rightarrow (ii) Let $x \in X$. If $y \in X$ is such that $y \neq x$ there exists disjoint $\text{pgr}\alpha$ -open sets U and V such that $x \in U$ and $y \in V$. Then $x \in U \subseteq X - V$ which implies $X - V$ is $\text{pgr}\alpha$ -open and $y \notin X - V$. Therefore $y \notin \text{pgr}\alpha\text{-cl}(U)$.

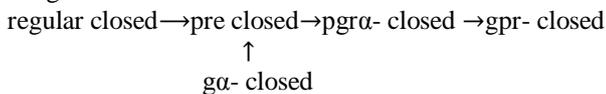
(ii) \Rightarrow (i) Let $x, y \in X$ and $x \neq y$. By (ii), there exists a $\text{pgr}\alpha$ -open U containing x such that $y \notin \text{pgr}\alpha\text{-cl}(U)$.

Therefore $y \in X - (\text{pgr}\alpha\text{-cl}(U))$. $X - (\text{pgr}\alpha\text{-cl}(U))$ is $\text{pgr}\alpha$ -open and $x \in X - (\text{pgr}\alpha\text{-cl}(U))$. Also $U \cap X - (\text{pgr}\alpha\text{-cl}(U)) = \emptyset$.

Hence (X, τ) is $\text{pgr}\alpha$ - T_2 space.

As application of $\text{pgr}\alpha$ -closed sets, four spaces namely, $\text{pgr}\alpha$ - $T_{1/2}$ spaces, $\text{pgr}\alpha$ - $T_{1/3}$ spaces, $\text{pgr}\alpha$ - T_b spaces, $\text{pgr}\alpha$ - $T_{3/4}$ spaces are introduced. The following implication diagram will be useful in this paper.

Diagram 3.14



Examples can be constructed to show that the reverse implications are not true. This motivates us to introduce the following spaces.

Definition 3.15 A space (X, τ) is called $\text{pgr}\alpha$ - $T_{1/3}$ if every gpr-closed set is $\text{pgr}\alpha$ -closed.

Definition 3.16 A space (X, τ) is called $\text{pgr}\alpha$ - T_b if every $\text{pgr}\alpha$ -closed set is regular closed.

Definition 3.17 A space (X, τ) is called $\text{pgr}\alpha$ - $T_{3/4}$ if every $\text{pgr}\alpha$ -closed set is $\text{g}\alpha$ -closed.

Theorem 3.18

(i) Every $\text{pgr}\alpha$ - T_b space is $\text{pgr}\alpha$ - $T_{1/2}$ space.

(ii) Every $\text{pgr}\alpha$ - T_b space is pre regular $T_{1/2}$ space.

(iii) Every pre regular $T_{1/2}$ space is $\text{pgr}\alpha$ - $T_{1/3}$ space.

(iv) Every $\text{pgr}\alpha$ - $T_{3/4}$ space is $\text{pgr}\alpha$ - $T_{1/2}$ space.

Proof: Straight forward.

The converse of the theorem need not be true as in the examples

Example 3.19 Let $X = \{a, b, c\}$ and $\tau = \{\emptyset, X, \{c\}, \{a, b\}\}$. Then (X, τ) is a $\text{pgr}\alpha$ - $T_{1/2}$ and pre regular- $T_{1/2}$ but not $\text{pgr}\alpha$ - T_b space.

Example 3.20 Let $X = \{a, b, c\}$ and $\tau = \{X, \emptyset, \{a, b\}\}$. Then (X, τ) is $\text{pgr}\alpha$ - $T_{1/3}$ but not pre-regular $T_{1/2}$ space.

Example 3.21 Let $X = \{a, b, c, d\}$ and $\tau = \{\emptyset, X, \{a, b\}, \{b, c, d\}\}$. Then (X, τ) is $\text{pgr}\alpha$ - $T_{1/2}$ but not $\text{pgr}\alpha$ - $T_{3/4}$ space.

Theorem 3.22 Let X be a $\text{pgr}\alpha$ - $T_{1/3}$ space. Then X is $\text{pgr}\alpha$ - $T_{1/2}$ if and only if it is pre regular $T_{1/2}$ space.

Proof: Suppose X is $\text{pgr}\alpha$ - $T_{1/2}$ and $\text{pgr}\alpha$ - $T_{1/3}$ space. Let A be gpr-closed set in X . then A is $\text{pgr}\alpha$ -closed set.

Since X is $\text{pgr}\alpha$ - $T_{1/2}$ space, then A is pre closed. Therefore X is pre regular $T_{1/2}$ space.

Conversely, we assume that X is pre regular $T_{1/2}$ space.

Suppose A is $\text{pgr}\alpha$ -closed set. Since every $\text{pgr}\alpha$ -closed set is gpr-closed set, and then A is gpr-closed set.

Since X is pre regular $T_{1/2}$ space then A is pre closed. This proves that X is $\text{pgr}\alpha$ - $T_{1/2}$ space.

Theorem 3.23

(i) If (X, τ) is an $\text{pgr}\alpha$ - $T_{1/3}$ space then for each $x \in X$, $\{x\}$ is either regular closed or $\text{pgr}\alpha$ -open.

(ii) If (X, τ) is an $\text{pgr}\alpha$ - T_b space then for each $x \in X$, $\{x\}$ is either regular closed or regular open.

(iii) If (X, τ) is an $\text{pgr}\alpha$ - $T_{3/4}$ space then for each $x \in X$, $\{x\}$ is either regular closed or $\text{g}\alpha$ -open.

Proof: Straight forward.

Theorem 3.24

(i) If X is $\text{pgr}\alpha$ - $T_{1/2}$ then every $\text{pgr}\alpha$ -continuous functions is pre continuous.

(ii) If X is $\text{pgr}\alpha$ - $T_{1/3}$ then every gpr-continuous function is $\text{pgr}\alpha$ -continuous.

(iii) If X is $\text{pgr}\alpha$ - T_b then every $\text{pgr}\alpha$ -continuous function is regular-continuous.

(iv) If X is $\text{pgr}\alpha$ - $T_{3/4}$ then every $\text{pgr}\alpha$ -continuous function is $\text{g}\alpha$ -continuous.

Proof: Straight forward.

Theorem 3.25 If X is pre-regular $T_{1/2}$ and $f: X \rightarrow Y$ then the following are equivalent

(i) f is gpr-continuous.

(ii) f is pre-continuous.

(iii) f is $\text{pgr}\alpha$ -continuous.

Proof: Suppose f is pgr -continuous. Let $A \subseteq Y$ be closed. Since f is pgr -continuous then $f^{-1}(A)$ is pgr -closed in X . Since X is pre regular $T_{1/2}$ space then $f^{-1}(A)$ is pre-closed.

Therefore f is pre-continuous. This proves (i) implies (ii).

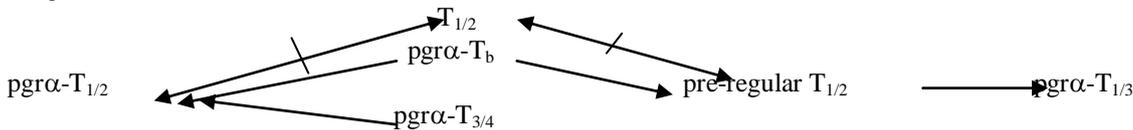
Suppose f is pre-continuous. Let $A \subseteq Y$ be closed. Since f is pre-continuous then

$f^{-1}(A)$ is pre-closed in X . We have $f^{-1}(A)$ is $\text{pgr}\alpha$ -closed. Therefore f is $\text{pgr}\alpha$ -continuous. This proves (ii) implies (iii).

Suppose f is $\text{pgr}\alpha$ -continuous. Let $A \subseteq Y$ be closed. Since f is $\text{pgr}\alpha$ -continuous then $f^{-1}(A)$ is $\text{pgr}\alpha$ -closed. Since every $\text{pgr}\alpha$ -closed set is pgr -closed then $f^{-1}(A)$ is pgr -closed. Therefore f is pgr -continuous. This proves (iii) implies (i).

$A \rightarrow B$ we mean A implies B but not conversely and $A \leftrightarrow B$ means A and B are independent of each other.

Diagram 3.26



IV. pgr \square -regular spaces and pgr \square -normal spaces

Definition 4.1 A topological space X is said to be an $\text{pgr}\alpha$ -regular space if for every $\text{pgr}\alpha$ -closed set F and each point x of X which is not in F , there exists disjoint pre-open sets U and V such that $x \in U$, $F \subseteq V$ and $U \cap V = \emptyset$.

Definition 4.2 A topological space X is said to be an $\text{pgr}\alpha$ -normal space if for every pair of disjoint $\text{pgr}\alpha$ -closed sets F_1 and F_2 in X , there exist disjoint pre-open sets U and V such that $F_1 \subseteq U$, $F_2 \subseteq V$, $U \cap V = \emptyset$.

Theorem 4.3 Let (X, τ) be a topological space. Then the following statements are equivalent

- (i) (X, τ) is $\text{pgr}\alpha$ -regular space.
- (ii) For each point $x \in X$ and for each $\text{pgr}\alpha$ -open neighborhood W of x , there exists a pre-open set of x such that $\text{pcl}(V) \subseteq W$.
- (iii) For each point $x \in X$ and for each $\text{pgr}\alpha$ -closed set not containing x , there exists a pre-open set V of x such that $\text{pcl}(V) \cap F = \emptyset$.

Proof: Let W be a $\text{pgr}\alpha$ -open neighborhood of x . Then there exists a $\text{pgr}\alpha$ -open set G such that $x \in G \subseteq W$. Since G^c is $\text{pgr}\alpha$ -closed set and $x \notin G^c$, by hypothesis there exists pre-open sets U and V such that $G^c \subseteq U$, $x \in V$ and $U \cap V = \emptyset$ and so $V \subseteq U^c$. Now $\text{pcl}(V) \subseteq \text{pcl}(U^c) = U^c$ and $G^c \subseteq U$ implies $U^c \subseteq G \subseteq W$. Therefore $\text{pcl}(V) \subseteq W$. Hence (i) implies (ii).

Let F be any $\text{pgr}\alpha$ -closed set and $x \notin F$. Then $x \in F^c$ and F^c is $\text{pgr}\alpha$ -open and so F^c is an $\text{pgr}\alpha$ -open neighborhood of x . By hypothesis, there exists a pre-open set V of x such that $x \in V$ and $\text{pcl}(V) \subseteq F^c$, which implies $F \subseteq (\text{pcl}(V))^c$. Then $(\text{pcl}(V))^c$ is a pre-open set containing F and $V \cap (\text{pcl}(V))^c = \emptyset$. Therefore X is $\text{pgr}\alpha$ -regular space. Hence (ii) implies (i).

Let $x \in X$ and F be an $\text{pgr}\alpha$ -closed set such that $x \notin F$. Then F^c is an $\text{pgr}\alpha$ -open neighborhood of x and by hypothesis there exists a pre-open set V of x such that $\text{pcl}(V) \subseteq F^c$ and therefore $\text{pcl}(V) \cap F = \emptyset$. Hence (ii) implies (iii).

Let $x \in X$ and W be an $\text{pgr}\alpha$ -open neighborhood of x . Then there exists an $\text{pgr}\alpha$ -open set G such that $x \in G \subseteq W$. Since G^c is $\text{pgr}\alpha$ -closed and $x \notin G^c$, by hypothesis there exists a pre-open set V of x such that $\text{pcl}(V) \cap G^c = \emptyset$. Therefore $\text{pcl}(V) \subseteq G \subseteq W$. Hence (iii) implies (ii).

Theorem 4.4 A topological space X is an $\text{pgr}\alpha$ -regular space if and only if given any $x \in X$ and any open set U of X there is $\text{pgr}\alpha$ -open set V such that $x \in V \subseteq \text{pgr}\alpha\text{-cl}(V) \subseteq U$.

Proof: Let U be an open set, $x \in U$.

So U^c is closed set such that $x \notin U^c$. Since X is a $\text{pgr}\alpha$ -regular space then there exist $\text{pgr}\alpha$ -open sets V_1 and V_2 such that $V_1 \cap V_2 = \emptyset$, $U^c \subseteq V_2$, $x \in V_1$. Since $V_1 \cap V_2 = \emptyset$, we have $\text{pgr}\alpha\text{-cl}(V_1) \subseteq \text{pgr}\alpha\text{-cl}(V_2^c) = V_2^c$.

Since $U^c \subseteq V_2$, we have $V_2^c \subseteq U$. Hence we have $x \in V_1 \subseteq \text{pgr}\alpha\text{-cl}(V_1) \subseteq V_2^c \subseteq U$.

Conversely, let F be a closed set in X and $x \in X - F$. So F^c is an open set such that $x \in F^c$.

Hence there exist a $\text{pgr}\alpha$ -open set U such that $x \in U \subseteq \text{pgr}\alpha\text{-cl}(U) \subseteq F^c$. Let $V = X - \text{pgr}\alpha\text{-cl}(U)$. So V is a $\text{pgr}\alpha$ -open set which contains F and $U \cap V = \emptyset$. Hence X is an $\text{pgr}\alpha$ -regular space.

Theorem 4.5 Let X and Y be topological spaces and Y is a regular space. If $f: X \rightarrow Y$ is closed, $\text{pgr}\alpha$ -irresolute and one to one then X is an $\text{pgr}\alpha$ -regular space.

Proof: Let F be closed set in X , $x \notin F$. Since f is closed mapping, then $f(F)$ is closed set in Y , $f(x) = y \notin f(F)$. But Y is $\text{pgr}\alpha$ -regular space then there are two $\text{pgr}\alpha$ -open sets U and V in Y such that $f(F) \subseteq V$, $y \in U$, $U \cap V = \emptyset$. Since f is $\text{pgr}\alpha$ -irresolute mapping and one to one so $f^{-1}(U)$, $f^{-1}(V)$ are two $\text{pgr}\alpha$ -open sets in X and $x \in f^{-1}(U)$, $F \subseteq f^{-1}(V)$, $f^{-1}(U) \cap f^{-1}(V) = \emptyset$. Hence X is $\text{pgr}\alpha$ -regular space.

Theorem 4.6 A topological space X is said to be an $\text{pgr}\alpha$ -normal space if and only if for every closed set F and for every open set G contain F there exists $\text{pgr}\alpha$ -open set U such that $F \subseteq U \subseteq \text{pgr}\alpha\text{-cl}(U) \subseteq G$.

Proof: Let F be a closed set in X and G be an open set in X such that $F \subseteq G$, G^c is a closed set and $G^c \cap F = \emptyset$. Since X is $\text{pgr}\alpha$ -normal space then there exist $\text{pgr}\alpha$ -open sets U and V of X such that $U \cap V = \emptyset$, $G^c \subseteq V$ and $F \subseteq U$, $U \subseteq V^c$.

We have $\text{pgr}\alpha\text{-cl}(U) \subseteq \text{pgr}\alpha\text{-cl}(V^c) = V^c$. Hence $F \subseteq U \subseteq \text{pgr}\alpha\text{-cl}(U) \subseteq V^c \subseteq G$.

Theorem 4.7 Let f be a closed and $\text{pgr}\alpha$ -irresolute mapping from a topological space X into a topological space Y . If Y is $\text{pgr}\alpha$ -normal, so is X .

Proof: Let F_1 and F_2 be closed sets in X such that $F_1 \cap F_2 = \phi$.

Since f is a closed map, we have $f(F_1)$, $f(F_2)$ are two closed sets in Y and $f(F_1) \cap f(F_2) = \phi$.

Since Y is $\text{pgr}\alpha$ - normal and f is $\text{pgr}\alpha$ - irresolute then there exists two $\text{pgr}\alpha$ - open sets U, V in Y such that $f(F_1) \subset U$, $f(F_2) \subset V$, $U \cap V = \phi$, also $f^{-1}(U)$, $f^{-1}(V)$ are $\text{pgr}\alpha$ - open sets in X and $F_1 \subset f^{-1}(U)$, $F_2 \subset f^{-1}(V)$, $f^{-1}(U) \cap f^{-1}(V) = \phi$. Hence X is $\text{pgr}\alpha$ - normal.

Theorem 4.8 Let X be a topological space. If X is a $\text{pgr}\alpha$ - regular and a T_1 space then X is an $\text{pgr}\alpha$ - T_2 space.

Proof: Suppose $x, y \in X$ such that $x \neq y$. Since X is T_1 - space then there is an open set U such that $x \in U$, $y \notin U$.

Since X is $\text{pgr}\alpha$ - regular space and U is an open set which contains x , then there is $\text{pgr}\alpha$ - open set V such that $x \in V \subset \text{pgr}\alpha\text{-cl}(V) \subseteq U$. Since $y \notin U$, hence $y \notin \text{pgr}\alpha\text{-cl}(V)$. Therefore $y \in X - (\text{pgr}\alpha\text{-cl}(V))$. Hence there are $\text{pgr}\alpha$ -open sets V and $X - (\text{pgr}\alpha\text{-cl}(V))$ such that $(X - (\text{pgr}\alpha\text{-cl}(V))) \cap V = \phi$. Hence X is $\text{pgr}\alpha$ - T_2 space.

References

- [1] K.Binoy Balan and C.Janaki, *On $\text{pgr}\alpha$ - closed sets in topological spaces*, *International Journal of Mathematical Archive*, 3(5), 2012, 2114-2121.
- [2] Y.Gnanambal, *On generalized pre-regular closed sets in topological spaces*, *Indian J. Pure Appl. Math.*, 28(3)(1997), 351-360.
- [3] E.D.Khalimsky, *Application of connected ordered topological spaces in topology*, *conference of Math. Departments of Povlsia*(1970)
- [4] E. Khalimsky, R.Kopperman, P.K.Meyer, *Computer graphics and connected topologies on finite ordered sets*, *Topology Appl.*, 36, No.1, (1990), 1-17.
- [5] T.Y.Kong, R.Kopperman, P.R.Meyer, *A topological approach to digital topology*, *Amer.Math.Monthly* 98 (1991), 901-917.
- [6] N.Levine, *Generalized closed sets in topology*, *Rend. Circ. Mat. Palermo*, 19(2)(1970), 89-96.
- [7] A.S.Mashhour, M.E.Abd El-Monsef and S.N El-Deeb, *On pre-continuous and weak pre-continuous*, *Proc.Math.Phys.Soc., Egypt*, 53(1982), 47-53.
- [8] O.Njastad, *On some classes of nearly open sets*, *Pacific J. Math.*, 15(1965), 961-970.
- [9] N.Palaniappan and K.C.Rao, *Regular generalized close sets*, *Kyungpook Math J.*, 33(1993), 211- 219.
- [10] M.Sheik John, *On w -closed sets in topology*, *Acta Ciencia Indica*, 4(2000), 389-392.

Differential Relay Reliability Impliment Enhancement of Power Transformer

Raju (M. Tech), K. Ramamohan Reddy, M.Tech (P.hd)

Department of EEE, KSRRM College of Engineering. Kadapa.

ABSTRACT: This paper presents an improvement of digital differential relay reliability for protecting a large power transfer is discussed. First, the Fourier sine and cosine coefficients required for fundamental, second third and fifth harmonics determination have been calculated using rectangular transfer technique. Then, these harmonics have been used in harmonics restrain and blocking techniques used in differential protection system. Simulation testes have been carried out on a variety of magnetizing conditions (normal aperiodic inrush and over excitation conditions) using Simulink? MATLAB.

Index-terms: Power transformer, Differential protection, Inrush current, Fourier coefficients, Rectangular transfer techniques.

I. INTRODUCTION

The main purpose of power systems is to generate, transmit, and distribute electric energy to customers without interruptions and in the most economical and safe manner. Power systems are divided into subsystems generation, transformation, transmission and distribution which are composed of costly components.

The role of protection ensures that, in the event of a fault, the faulted element must be disconnected from the system for isolating the fault to prevent further damage to the components of the system through which the fault currents were flowing. A power transformer is mostly protected against internal fault using a differential protection which is sensitive and a fast clearing technique. This technique of protection detects nonzero differential current, then activates a circuit breaker that disconnects the transformer. However, this nonzero differential current may be produced by transformer magnetization, due to so called inrush current of over-excitation, and may cause the protective system to operate unnecessarily. This magnetization current is a transient current that appears only when a transformer is first energized or after clearing external faults.

During periodic inrush condition due to over-excitation the third and fifth harmonic components are largely seen: however, during the normal aperiodic inrush conditions, the second harmonic is relatively high.

The transformer differential protection scheme has to be improved so that it can distinguish between nonzero differential current produced by magnetization current and that produced by internal fault. Several methods have been proposed to blind the differential protection system to magnetization current where the harmonic components have been used as means of detection. However, the digital computer based protection offers a number of advantages over the conventional ones. So, the security and reliability have been improved; it remains only to develop an efficient algorithm requiring less time consuming calculations.

The alternative approaches to the digital protection of power transformer have been proposed to date; one using a digital filtering approach and the other using sine and cosine wave correlation to yield the fundamental and higher harmonic components required for protection. This paper presents a new approach which the sine and cosine Fourier coefficients are expressed in terms of rectangular transfer coefficients that are obtained from the data samples by only additions and subtractions.

II. TRANSFORMER DIFFERENTIAL PROTECTION

The most important devices employed in the protection system are protective relays. These devices may be flexible, economic and provide reliable, fast and inexpensive protection. The IEEE standard defines a protective relay as "a relay whose function is to detect faults or other power conditions of an abnormal or dangerous nature and to initiate.

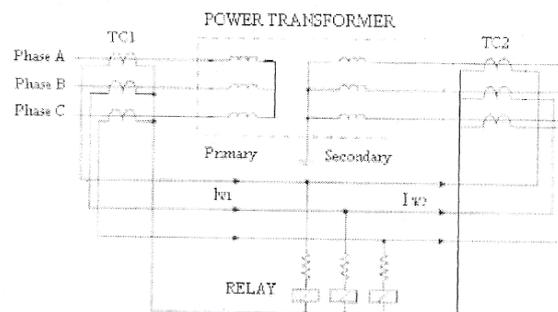


Fig.1 Typical Differential power transformer protection relay.

Appropriate control circuit action the differential protection principle is simple and provides the best protection for the phase and ground faults.

Differential relay is generally used for protection the power transformer against internal fault. Figure I shows a typical differential relay connection diagram.

Even differential protection is relatively simple to apply, but it has problems. One of the problem of the differential relay is its operation due to transformer magnetizing current which is well known, this current appears on only one input to the differential relay (from the side of energization), thus the relay sees this situation as an internal fault. Figure 2 illustrates the typical current waveform present during a one phase transformer bank energization.

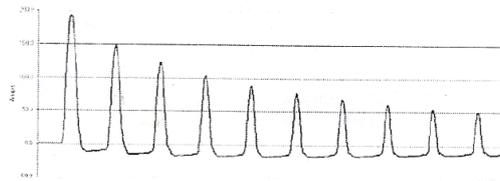


Fig.2 Transformer Inrush (one Phase)

An inrush current is the surge of transient current that appears in a transformer. The exciting voltage applied to the primary of the transformer forces the flux to build up to a maximum theoretical value of double the steady state flux plus remanence, therefore the transformer is greatly saturated and draws more current which can be in excess of the full load rating of the transformer windings.

This current is high magnitude, harmonic-rich currents generated when transformer cores are driven into saturation.

$$\phi_{MAX} = 2\phi_M + \phi_R \quad (1)$$

Although it is usually considered as a result of energizing a transformer, the magnetizing inrush may be also caused by.

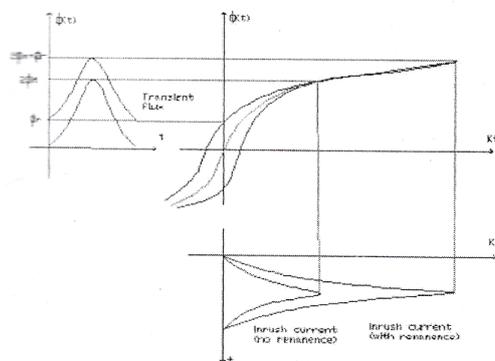


Fig.3 Typical curve of fluxes and inrush current.

1. Occurrence of an enternal fault,
2. Voltage recovery after clearing an external fulut,
3. Change of the type of a fault,
4. Energizing of a transformer in parallel with a transformer that is already in service.

The solpe, magnitude and duration of inrush current depend on several factors.

- Size of a transformer
- Impedance of the system from which a transformer is energized.
- Magnetic properties of the core material,
- Magnetic residual in the core,
- Why a transformer is switched on (inner, outer winding, type of switchgear)
- When a transformer is switched on.

A. Problems Caused by Inrush Current

An important feature of this inrush current is that the current is not pure fundamental frequency waveform. From a power quality point of view, the magnetizing inrush current can be considered as a distorted wave with two kinds of disturbances.

Harmonics: Part research has shown that magnetizing inrush produces currents with a high second harmonic content (8), with relatively low third harmonic content (9) and higher harmonics with different small values, so that can be neglected.

Unblance: Current unbalance cannot be considered a disturbance. Asymmetrical load produce unbalance currents. In the same way, the magnetizing inrush current produces current unbalance during magnetization, but is not a fault, and the differential relay must not trip.

Other disturbances caused by inrush may occur due to:

1. Incorrect operation and failures of electrical machines and relay systems.
2. Irregular voltage distribution long the transformer windings.
3. High amount of voltage drop at the power system at energization timers
4. Electrical and mechanical vibrations among the windings of the transformer.

B. Differential Protection Methods.

The most important means of protection based on the comparison of the transformer primary and secondary currents. When these currents deviate from a predefined relationship, an internal fault is considered and the transformer is de-energized. However, during transient primary magnetizing inrush conditions, the transformer can carry very high primary current and no secondary current.

1. Power differential method: this method is based on the idea that the average power drawn by a power transformer is almost zero on inrush, while during a fault the average power is significantly higher
2. Rectifier relay: this method is based on the fact that magnetizing inrush current is in effect a half-frequency wave. Relays based on this method use rectifiers and have one element functioning on positive current and one on negative current.
3. Waveform recognition: it is the method of measuring dwell-time” of the current waveform, that is, how long it stays close to zero, indicating a full dc-offset, which uses to declare an inrush condition. Such relays typically expect the dwell time to be at least ¼ of a cycle, and will restrain tripping if this is measured.
4. Flux-current: A new simple and efficient technique for inrush current reduction based on the calculated flux in the core. As its advantage, this approach tides together the cause of the problem (saturation of the core as a source of the current unbalance) with the phenomenon used for recognition.
5. Cross blocking: it is a “method that blocks all tripping if any relay detects inrush. Any of the relays that use single-phase inrush detection methods can utilize cross blocking.
6. Harmonic current restraint: This is the most common method and widely used for the detection of inrush current in power transformer.

C. Harmonic Current Restraint

Different schemes currently used to distinguish between magnetizing Inrush and fault current are based on:

1. Second harmonics restraint principle.
2. Voltage restraint principle.
3. Restraint principle based on currents and voltages of the transformers.

Simple 2nd harmonic restraint: This method has been used for many years and simple employs a percentage level of 2nd harmonic content (or THE in some relays) in the differential current. If the 2nd harmonic content present in the waveform is above a threshold (typical thresholds are between 15 and 35% of fundamental) the relay is restrained. This is simply a per-phase calculation of 2nd harmonic current (in Amps) divided by fundamental current (in Amps) .

Shared 2nd harmonic restraint: The same methods as described above with the exception that the numerator is the sum of the 2nd harmonic current from all three differential currents is 9A and the particular phase of interest (this calculation is performed for each phase) has 10A of fundamental its restraining quantity is 90%.

III. MAGNETIZING CURRENT ALGORITHM

In a large power transformer, any switching action can produce a large current peak due to the saturation of the transformer iron core. Owing to this core saturation, the inrush current contains, in addition to the harmonic components, a decaying dc current. Therefore, the inrush current can be modeled as follows:

$$i(t) = I_o \exp(-\lambda t) + \sum_{k=1}^n I_k \sin(k\omega_1 t + \theta_k) \quad (2)$$

Where k determines the order of harmonic, and ω_1 is the frequency of the fundamental component. The decaying dc current can be represented by a Taylor expansion of two terms:

$$I_o \exp(-\lambda t) \approx I_o - I_o \lambda t \quad (3)$$

If it is assumed that the inrush current does not contain more than five harmonics, Eq.(1) becomes.

$$i(t) = I_o - I_o \lambda t + \sum_{k=1}^5 I_k \cos \theta_k \sin(k\omega t) \quad (4)$$

$$S_5 = \left(1 + \frac{1}{16}\right) S_3'$$

$$C_5 = \left(1 - \frac{1}{16}\right) C_5'$$

The harmonics components are found to be

$$I_1 = \frac{2}{12} [S_1^2 + C_1^2]^2 \tag{13a}$$

$$I_2 = \frac{2}{12} [S_2^2 + C_2^2]^2 \tag{13b}$$

$$I_5 = \frac{2}{12} [S_5^2 + C_5^2]^2 \tag{13c}$$

After extraction of the fundamental, the second and the fifth harmonic components, these harmonic components will be used to produce restraining signal that may be used to block there relay. Otherwise, for internal fault case, the relay operates.

IV. PROTECTION SYSTEM IMPLEMENTATION

The Protection System approach has been implemented using Matlab/Simulink with the necessary tool box. The Matlab is powerful software program used for any test and simulation. The characteristics of the differential protection scheme that has been used are plotted in Where there are two straight lines give with a slope of K1=0.25 and a slope of K2=0.6 which range from Irt0 to Irt1 and from Irt1 to Irt2, respectively, and a horizontal, straight line defining the relay minimum pickup current, Iop0=0.3A. The relay operating is located above the slope, and the restraining region is below the slope.

A dual-slope percentage characteristic provides further security for external faults. It is represented as a dashed line in.

The dual-slope percentage pattern adds a restraint area and avoids mal-operation caused by saturation. In comparison with a single-slope percentage scheme, the dual-slop percentage current differential protection can be regarded as a better curve fitting of transformer operational principles.

A Systems Description And scopes

The Simulink model a illustrated in consists of a three-phase transformer rated 225 kVA, 2400 V/600V, 60Hz, connected to a 1 MVA, 2400 V power network. A 112.5kW resistive load (50% of transformer nominal power) is connected on the 600V side. Each phase of the transformer consists of two windings both connected in wye with a grounded neutral. In a system relaying block, the currents that have been measured on Buses B1 and B2 pass through a second order Butterworth low pass filter with a cut

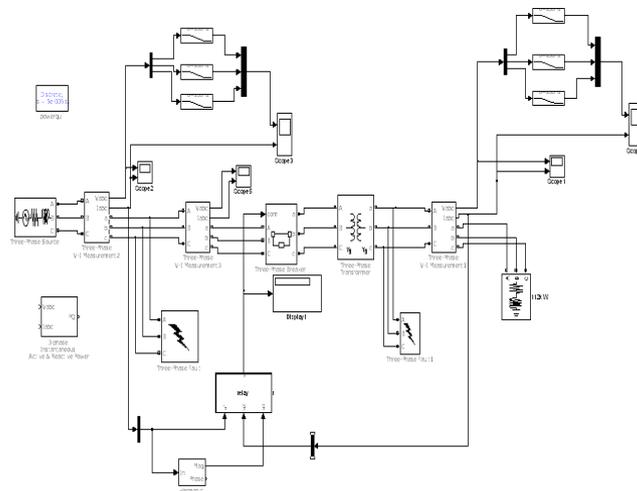
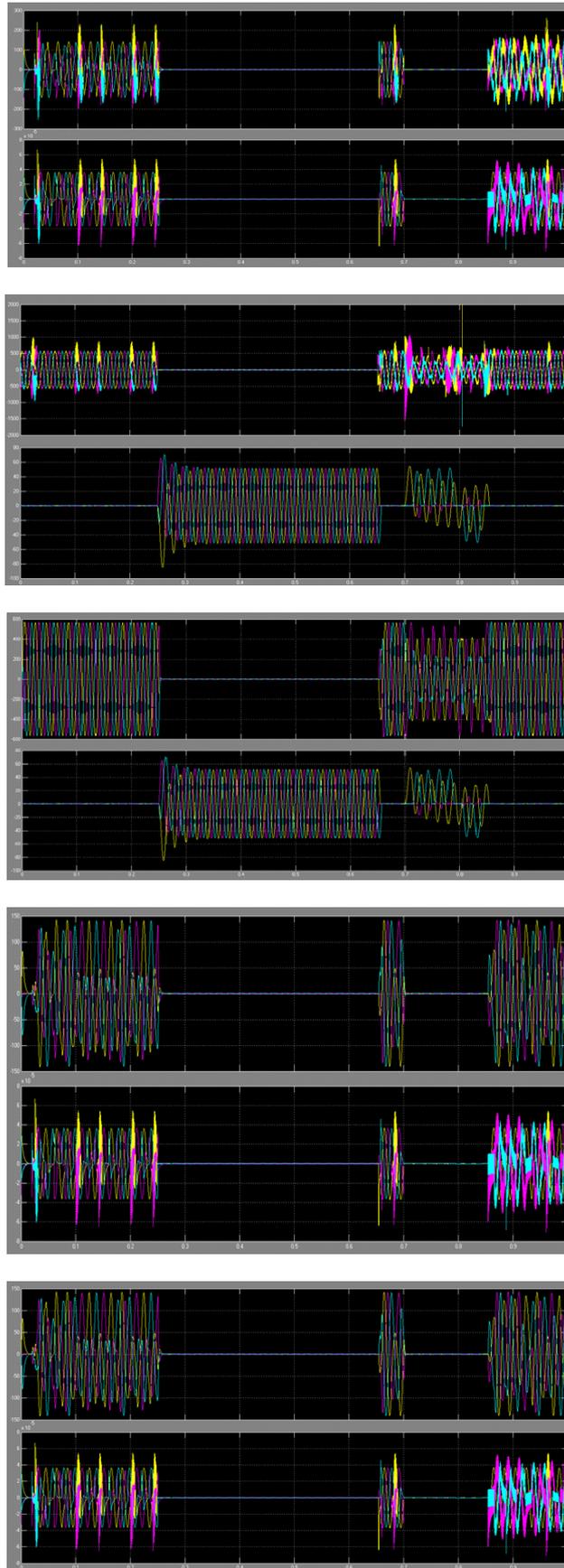


Fig.simulation block diagram



Out puts of simulation results

Off frequency of 600 Hz, which offers a maximum flat response in the pass band and a quite good attenuation slope After that, the differential and restrain currents using blocks included in Simulink library and our algorithm, have been calculated. The generated signals are used in the relay operational principles.

V. SIMULATION RESULTS

The was simulated in MATLAB using the Simpower system toolbox of SIMULINK. The digital by simulation for magnetization currents and internal fault cases. These currents are generated when the circuit breaker is closed to connect the transformer and external fault appears as shown in Fig 6. the currents are measured by current transformers on buses B1 and B2 and then introduced to the relay. Some parameters have been made variable to allow performing all possible cases of test. Two test cases have been performe.

a) switching on the transformer and then applying an external fault as shown in Fig6

b) switching on the transformer and then applying an internal fault as shown in Fig7

figure 6 shows the plots of the differential currents then the transformer is switched on at $t=0.08$ sec and then an external fault at 0.25 sec and finally this fault cleared at 0.65 sec. In fig(7) the differential current as well the restrain current are shown for case (b) switching on the transformer at $t=0$ and then applying an internal fault at $t=0.6$ sec. However, fig.8 shows the plots of test case(b) for the relay trips. The output and response time of the relay are shown in this figure. However, the trip times that have been found. Include the waiting time of one cycle of the power frequency. This delay has been introduced to prevent false trip conditions. It is possible to reduce the time delay to achieve faster tripping. It can be noted that the relay exhibits a good response in all considered cases. This method allows obtaining a rapid and accurate response of the digital protection scheme. Moreover, it provides a good discrimination between the inrush current and internal fault current

VI. CONCLUSION

This paper, presents, an attempt has been made through the use of MATLAB/SIMULINK to test a new approach applied to digital differential protection relay for a large power transformer. First, the Fourier sine and cosine coefficients required for fundamental, second, third and fifth harmonics extraction have been calculated using rectangular transfer technique. Then, these harmonic components have been used in harmonics restrain and blocking techniques which may be utilized in differential protection system. Testes have been carried out on a variety of magnetizing conditions (normal aperiodic inrush and over excitation conditions due to external fault) as well as internal fault. It can be noted that, from the obtained simulation results using Simulink/MATLAB, the developed scheme provides good discrimination between the magnetizing current and the internal fault current.

REFERENCES

- 1) J.S thorp, A.G.Phadke, A Microprocessor Based three-Phase Transformer Differential Relay, IEEE Trans, PAS-101,P426,1982
- 2) M.A.Rahman, P.K.Dahs, "Fast Algorithm for Digital Protection of Power Transformer" IEEE proc., Vol 129-C2,P,79,1982.
- 3) M.A.Rahman, P.K.Dahs, "Fast Algorithm for Digital Protection of Power Transformer" IEEE proc., Vol 129-C2,P,79,1982.
- 4) B Kaszenny, and A.kuliidjiant, "An Improved Transformer inrush restraint current Algorithm" 53rd Annual conference for protective relay engineering, College, April 11-13,2000.
- 5) F.Meclie, R.Girgis,Z. Gajic, Power Transformer Characteristics and their Effects on protective Relays" 33rd Western protective Relays conference, October 17-19, 2006.
- 6) R. Bouderbala, H.Bentarzi and O. Abderrahmane. "A New approach Applied to Digital Differential Protection for a Large Power Transformer" the 9th WSEAS international conference on Recent Researches in Circuits, Systems, Electronics, Control & Signal Processing Proceeding, PP.202-205, ISBN:978-960-474-262-2, Athena, Greece, Dec.29-31-2010.
- 7) A.Ouadi, H.Bentarzi, and J.C.Maun, "Improvement of phasor Measurement Unit Performance" the 9th WSEAS international control & signal processing proceeding, pp.202-205, ISBN: 978-960-474-262-2, Athena, Greece, Dec.29-31-2010

A Novel Family of Forward Converter with Energy Regenerative Snubber

S. John Prakesh¹, D. Venkateswara Rao²

¹(Associated professor, Department of EEE, Nova institute of Technology, India,
²(Student, Department of Electrical & Electronics Engineering, KL University, India,

ABSTRACT: An energy regenerating snubber for a forward converter is presented. The proposed snubber uses a tertiary transformer winding and beneficially exploits transformers' leakage inductances. The proposed snubber is capable of resetting the transformer, as well as eliminating leakage inductance voltage spike across the power switch. The proposed snubber recycles the recovered energy partially to the load and partially back to the source. This paper presents theoretical analysis, simulation and experimental results.

Keywords: DC-DC power conversion, Duty Cycle Limitation, PWM, RCD Clamp, snubbers, ZVS

I. INTRODUCTION

THE forward converter is a relatively simple and popular topology and retains many features of the buck converter. The greatest advantage of the forward converter is its ability to provide multiple outputs by means of an isolation transformer. With a proper choice of the transformers' turn-ratio the forward converter can attain wide voltage step-down useful for offline applications. In addition, the forward converter provides the load with non pulsating output current and, therefore, is well suited for high output current applications. Moreover, with a linear small-signal control-to-output transfer function, a fixed output filter resonant frequency and no right half-plane zero the forward converter is quite easy to control. These advantages make the forward converter the designers' choice for low to medium isolated offline power applications.

The transformer, however, brings in two main problems. First, the transformer core needs to be reset to prevent the magnetizing current build up and transformer Saturation Second, the transformer leakage inductance has to be discharged to prevent voltage spike and switch over-voltage breakdown. These problems were tackled by many researchers in the past and still are of interest at present.

As compare to the non dissipative snubber, the proposed energy regenerative snubber has all the advantages of the former, however, eliminates the need for discrete inductor. When properly designed, the proposed snubber : 1) allows zero-voltage turn OFF; 2) provides transformer reset; 3) absorbs the leakage energy; 4) controls switch dv/dt ; 5)

In the classical configuration, the forward converter uses a tertiary winding for core reset [1], [2] as shown in Fig. 1(a). By this method, the magnetizing inductance energy is recycled to the source. The reset winding somewhat complicates the transformer structure, and a reset diode is also required. However, a major disadvantage of the reset winding method is that the transformer leakage inductance discharge spike cannot be put OFF. The leakage spike sums up with the voltage induced by the reset winding and causes a high voltage stress across the active switch. Another disadvantage is the hard turn-off of the active switch.

An alternative method of forward converter transformer reset is using a resistor, capacitor, and diode (RCD) clamp [3]. The RCD clamp circuit is quite simple, consisting only of three components, a resistor, a capacitor, and a diode, and requires a simpler two winding transformer as shown in Fig. 1(b). The clamp absorbs the magnetizing inductor energy as well as provides a discharge path for the leakage inductance. Thus, with RCD clamp, the transformer is totally reset and the power switch experiences significantly less voltage stress than that with the reset-winding method. For these reasons, RCD clamp forward converter is widely used in power supply industry. The main disadvantage of RCD clamp is that the recovered energy is dissipated. and the overall converters' efficiency deteriorates. Moreover, with the RCD clamp, the switch still experiences a hard turn-off.

The two switch forward converter is a viable topology. A variation of the non dissipative snubber for the two switch forward converter is suggested. Interesting transistor forward converter with inherent reset mechanism was also proposed.

In this paper, a regenerative snubber for the single switch forward converter is proposed. The suggested snubber is an evolution of the reset and snubber circuits of Figs. 1(a) and (b) and attains the best of the

transfers the recovered energy to both to the source and to the load side; and 6) provides zero-current turn ON. This paper presents the principle of operation of the proposed regenerative snubber, as well as simulation and experimental results.

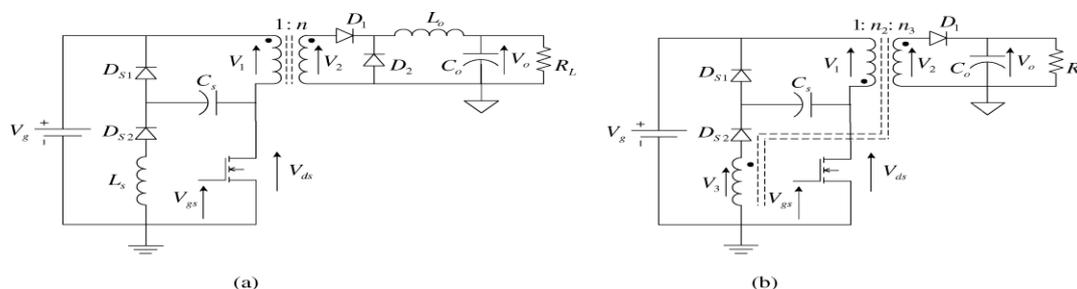


Fig.1. (a) Forward converter with the non dissipating snubber. (b) Fly back converter with the regenerative snubber

II. PROPOSED TOPOLOGY

A forward converter with the proposed energy regenerative snubber. The snubber is comprised of the capacitor C_1 , tertiary transformer winding n_3 , and a pair of snubber diodes, $DS1$ and $DS2$. The resemblance with the regenerative snubber of the fly back converter in Fig. 1(b) is obvious. However, the forward converter operation is very much different from the fly back converter operation. The dissimilarities arise due to the different winding polarity arrangement, reset requirement of the transformer core and nature of the rectifier, and the output filter. In a fly back converter, the input and output circuits conduct consecutively whereas, in forward converter, the primary and secondary conduct simultaneously. While the capacitive filter in fly back converter clamps the secondary winding to the output voltage, inductive output filter in the forward converter can sustain primary current and affects the

snubber. Fly back converter operates by charging the transformer core and then discharging the magnetizing energy to the load, thus, transformer reset is natural whereas, a forward converter lacks such reset mechanism. Accordingly, the snubber task in a fly back converter is to absorb the leakage energy whereas, in the forward converter the snubber should provide both core reset and discharge of leakage inductance.

III. ENERGY REGENERATIVE SNUBBER

The transformer is central to analyzing the operation of the forward converter with the proposed regenerative snubber as shown in Fig. 2. The analysis to follow relies on the three winding transformer model. Here, the transformers' magnetizing and the leakage inductances are all referred to the primary.

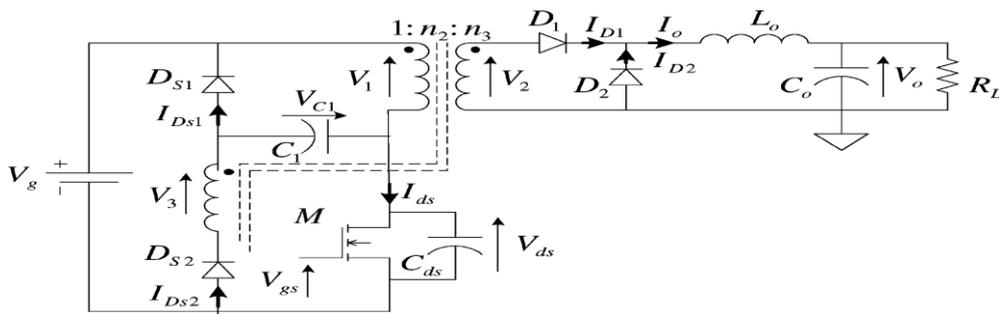


Fig. 2. Forward converter with the proposed energy regenerative snubber

IV. SIMULATION DIAGRAM AND RESULTS

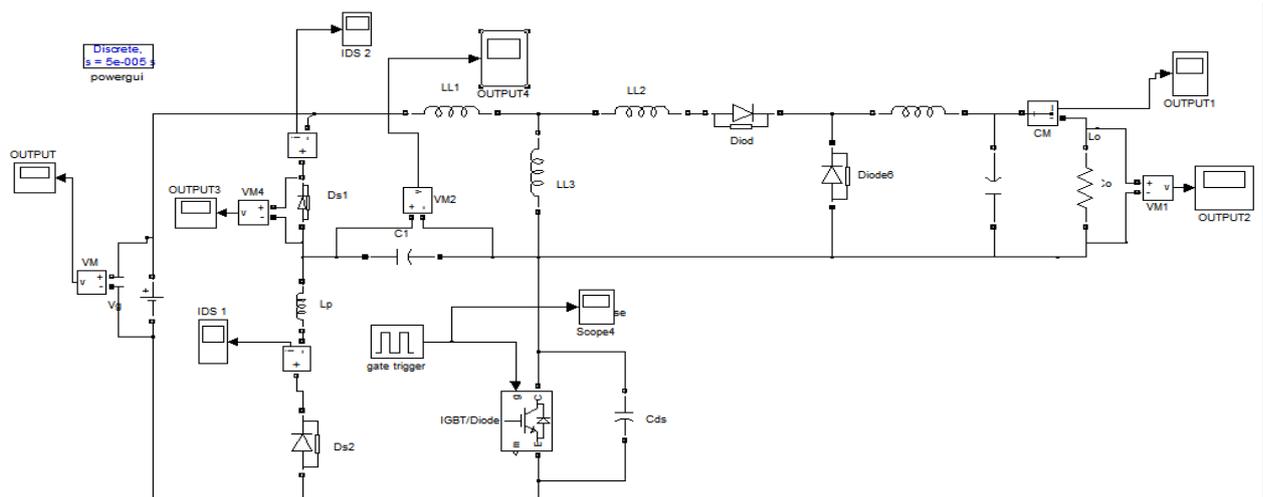


Fig-3 forward converter with snubber by using buck converter

An experimental forward converter with energy regenerating snubber was built and tested. The converter specs were as follows. Input voltage: $V_g = 28\text{Vdc}$; output voltage: $V_o = 5\text{Vdc}$, output current: $I_o \text{ max} = 5\text{A}$; switching frequency: $f_s = 200 \text{ kHz}$. Through the experiments, the converter was operated in a wide duty cycle range: $D_{\text{max}} = 0.7$, $D_{\text{min}} = 0.1$. The transformers' secondary and tertiary turns ratio were: $n_2 = n_3 = 0.4$. The primary inductance was $L_p = 411 \mu\text{H}$ and the leakage inductances were: $LL1 = 20$, $LL2 = 6$, and $LL3 = 6 \mu\text{H}$, respectively. The snubber capacitance was found using as:

$C_1 \approx 3\text{nF}$. The output filter was comprised of $L_o = 260 \mu\text{H}$ and $C_o \approx 100 \mu\text{F}$, respectively.

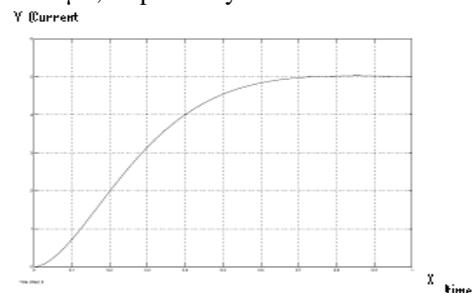


Figure 3.1 current waveform

In a typical nonintegrated switching dc–dc converter, significant energy is dissipated by the parasitic impedances of the interconnect among the nonintegrated devices (the filter inductor, filter capacitor, power transistors, and pulse width modulation circuitry) Moreover, the integrated active devices of a pulse width modulation circuit are typically fabricated in an old technology with poor parasitic impedance characteristics. Integrating a dc–dc converter with a microprocessor can potentially lower the parasitic losses as the interconnect between (and within) the dc–dc converter and the microprocessor is reduced. Additional energy savings can be realized by utilizing advanced deep sub micrometer fabrication technologies with lower parasitic impedances. The efficiency attainable with a monolithic dc–dc converter, therefore, is higher than nonintegrated dc–dc converter.

Generally, in order to attain higher efficiency, buck derived converters should be operated at as large a duty cycle as possible. The duty cycle, however, is limited by switching stress. To find an optimum operating point, switch utilization function is considered next. As you can see there are only four main components: switching power MOSFET Q1, flywheel diode D1, inductor L and output filter capacitor C1. A control circuit (often a single IC) monitors the output voltage, and maintains it at the desired level by switching Q1 on is and off at a fixed rate (the converters operating frequency), but with a varying duty cycle (the proportion of each switching period that Q1 turned on).When Q1 is turned on, current begins flowing from the input source through Q1 and L, and then into C1 and the load. The magnetic field in L therefore builds up, storing energy in the inductor . with the voltage drop across L opposing or .bucking. part of the input voltage. Then when Q1 is turned off, the inductor opposes any drop in current by suddenly reversing its EMF, and now supplies current to the load itself via D1.

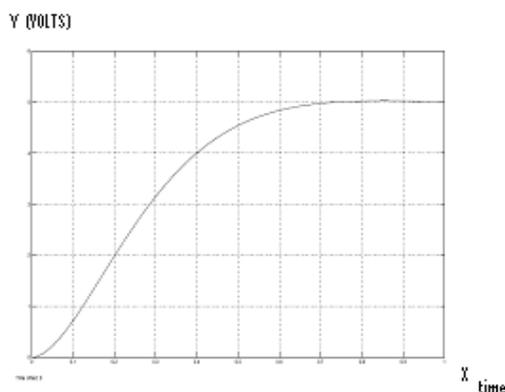


Figure 3.2: voltage waveform

V. CONCLUSION

The non dissipating snubber reported in uses a discrete inductor to perform the reversal of the snubber capacitor voltage and feed the energy back to the source.

Recycling of the energy takes place during the clamp state. However, when the converter operates with low duty cycle, the magnetizing energy the capacitor absorbs may be insufficient to raise the capacitor voltage to the value of the source voltage. Consequently, the clamp state will not occur and the non dissipating snubber changes the operating mode. The undercharged snubber capacitor cannot provide true ZVS condition for the power switch and the power stage efficiency deteriorates The non dissipating snubber provides no means to make any readjustments in order to improve the performance. Contrary to that, the proposed regenerative snubber uses a tertiary transformer winding and beneficially exploits the transformer leakage inductances. The tertiary winding turn-ratio introduces another degree of freedom and can be adjusted to ensure a complete recharge of the snubber capacitor to the full value of the source voltage. As a result, mode changes can be avoided and perfect ZVS conditions for the power switch can be provided for a wide range of operation conditions. This is one of the most important advantages of the proposed snubber circuit over the previously reported counterpart. The disadvantages of the proposed snubber are the more elaborated transformer structure, and second, the voltage spike which appears across the free-wheeling diode during the interval C.

ACKNOWLEDGEMENTS

The proposed snubber is capable of resetting the transformer, as well as eliminating leakage inductance voltage spike across the power switch. The proposed snubber recycles the recovered energy partially to the load and partially back to the source.

REFERENCES

- [1] R.W.Ericson and D. Maksimovic, *Fundamentals of Power Electronics* 2nd ed. Norwell, MA: Kluwer, 2000.
- [2] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics; Converter, Applications and Design*, 3rd ed. New York: Wiley, 1989.
- [3] C. D. Bridge, “Clamp voltage analysis for RCD forward converters,” in *Proc. IEEE APEC*, 2000, pp. 959–965.
- [4] M. Domb and R. Redl, “Nondissipative turnoff Snubber in a forward converter: Analysis design procedure, and experimental verification,” in *Proc. PCI*, Oct.1985, pp. 54–68.
- [5] T. Ninomiya, T. Tanaka, and K. Harada, “Analysis and M.Jinno” efficiency improvement for SR forward converter switch lc snubber,” *IEEE Trans. Power electron* .Vol.16, no.6, pp.812-820, nov.2001



S. John prakesh received the B.Tech Degree in Electrical And Electronics Engineering from JNT University Hyd,India 2009.The M.Tech (PEPS) Power Electronics And Power Systems at KL University,Vijayawada,India 2011.present i am working as Hod of EEE Department in Nova Institute of Technology ,ellore, India



D.Venkateswara Rao Received the B. Tech Degree in Electrical And Electronics Engineering from JNT University Hyd,India 2008. The M.Tech (PEPS) Power Electronics And Power Systems at KL University,Vijayawada,India 2011.Present I am working Research area at Solar power generation

Effect of Air Velocity, Fuel Rate and Moisture Content on the Performance of Updraft Biomass Gasifier using Fluent Tool

Hina Beohar^a, Bhupendra Gupta^{a*}, Dr. V. K. Sethi^b, Dr. Mukesh Pandey^b,
Hemant Parmar^c

^aJabalpur Engineering College, Jabalpur, M.P.

^bUIT, RGPV, Bhopal, M.P.

^cUjjain Engineering College, Ujjain, M.P.

Abstract: The use of biomass gasification for conversion of hydrocarbons to permanent fuel gas mainly composed of H₂, CO, CO₂ and CH₄. In this study the performance of Updraft gasifier is carried out with ANSYS FLUENT software, to study the effect of various operating parameters such as air velocity, fuel rate and moisture content on the performance of updraft gasifier are discussed. It was found incredible effect to varying the different operating parameters in different range. With increasing air velocity and moisture content, the value of H₂ first increased then decreased but CO initially decreased then increased, CO₂ increased gradually. When increasing fuel rate, the value of H₂ and CO₂ first decreased then increased, and the value of CO increased then decreased. So it was found the optimum value of air velocity, fuel rate and moisture content is 5.4 m/s, 8.8 kg/h and 25% respectively.

Key words: Biomass, up-draft gasifier, air velocity, fuel rate and moisture content, gascomposition, FLUENT simulation.

I. Introduction

In India, biomass has always been an important energy source. Although the energy scenario in India today indicates a growing dependence on the conventional forms of energy, about 32% of the total primary energy use in the country is still derived from biomass and more than 70% of the country's population depends upon it for its energy needs. Our current work focused on the computational modeling of updraft gasifier which using the long stick wood as a biomass material. CATIA V5R16 software used to construct the model and ANSYS FLUENT was employed to analysis the various parameters of updraft gasifier system.

Nomenclature

A_g cross section area of gasifier (mm)	V_a velocity of air flow (m/s)
D typical dimension (mm)	C species concentration (%)
F_m moisture content (%)	P_g producer gas
F_{O_2} fraction of oxygen (%)	E activation energy (J/mol)
L length of the wood (mm)	L_c length of the charcoal gasification (mm)
m mass flow rate of fuel (kg/h)	Greek symbols
T temperature (K)	α mass fraction
t_c char gasification time (s)	ϵ_i volume fraction of phase
V volume of the wood (m ³)	μ viscosity, kg/ms
V_f velocity of fuel flow (m/s)	ρ density, kg/m ³

II. Literature Review:

Marta Muilenburg et al [7] focused on the computational modeling of downdraft gasifier bed which will be implemented by the University of Iowa Oakdale power plant. Gambit software was used to construct the meshes and ANSYS FLUENT was employed to test the various factors so as to complete the modeling of the down draft gasifier system for this project.

S. Murgia et al [11] presented a comprehensive CFD model of an air blown up-draft gasifier has been setted up and tested using the CFD code MFiX. This paper the performance of gasifier characterize by two parameters such as Temperature and Gas composition in the function of time.

Avdhesh Kr. Sharma [3] Presented 75 kWth, downdraft gasifier system has been carried out for obtaining temperature profile, gas composition, calorific value and trend for pressure drop across the porous gasifier bed. In firing mode, the higher temperature in bed tends to better conversion of non-combustibles component (like CO₂, H₂O) into combustibles component (like CO, H₂) in the resulting gas and thus improves in calorific value of product gas.

Weihong Yang et al [13] Present work an experimental fixed bed gasifier is utilized to investigate the gasification of biomass using higher temperature air up to 1473K. A model has in bed been formulated for the prediction of the main chemical and physical processes and used to study the influence of oxygen concentration, air flow rate, gas composition. The temperature of feed gas increased a higher gasification rate and higher molar fraction of feed gases (CO, H₂). Increased oxygen concentration leads to higher peak values of the fuel gas concentrations, a higher gasification rate.

P. Abdul Salam *et al* [8] paper presents the experimental study on air gasification of charcoal biomass gasifier. Performance of a gasifier depends on the design of the gasifier, type of fuel used and air flow rate. The experiments were carried for an equivalence ratio of 0.25. The effect of air velocity and type of the distributor on the gasification performance was discussed. The concentration of CO₂ (12.4–15.4 vol %) is found to show increasing trend with increase in velocity of air. The concentration of CH₄ was low in the range of 0.51–0.94 vol% and remained almost constant with the increase of air velocity.

A. Saravanakumar *et al* [1] present work focus of made on development of a gasifier using long sticks of wood as feed materials. And the performance of gasifier is characterized in terms of air flow rate, fuel flow rate, gas composition. When increased air flow rate the higher rate O₂ oxide that caused composition H₂ increased and CO decreased. When fuel flow rate increased the biomass consumption rate is decreased that reasons H₂ decreased.

A. Saravanakumar *et al.* [2] this concept a 25m³/hr capacity gasifier was design and constructed. Since this gasifier attains a high-energy release rate per unit area due to high inlet air velocity and activated reduction in the combustion zone. In all zones, the heat balance equations show a good fit between the theoretical and experimental value. The gasifier was operated in a batch mode, both bottom-lit and top-lit, and air flow and gas out flow were measured.

D.F Fletcher *et al* [5] content the description of CFD model to simulate flow and reaction an entrained flow biomass gasifier. ANSYS CFX used to provide detail information of gas composition and temperature at the outlet of gasifier.

M. Miltner *et al* [6] this paper show the joint application of process simulation and CFD. And modeling approach is essentially focused on treatment of heterogeneous combustion and prediction of gases emission such as carbon monoxide and nitrogen oxide.

PengmeiLv *et al* [9] This study applies a self heated gasifier as the reactor and uses char as catalysis to study the characteristics of hydrogen production from biomass gasification. Air and oxygen/steam are utilized as gasification agents. The experimental result indicate that compared to biomass air gasification, oxygen/steam gasification improve hydrogen yield depending on the volume of downdraft gasifier, and also nearly doubles the heating value of fuel gas.

III. Modeling of Gasifier

The work includes three parts. In first part we create model using CATIA V5R16 software. In second part put all input parameters for simulation and third part simulation done and results analyzed by the FLUENT ANSYS software.

(i) Design of gasifier:

In Present work long stick wood up-draft gasifier was constructed. The gasifier in fig.1 was design and construct using mild steel material. It is rectangular in cross-section. the total height of gasifier is 1.34 m. and total length and breadth of gasifier is 75cm and 60cm respectively. The hopper is the primary storage area for the fuel wood. Long sticks wood packed upto the full capacity of the hopper. The air enters, air nozzle its diameter is 1.5 inch. There are two air nozzle fixed with adjacent side of hearth of gasifier.

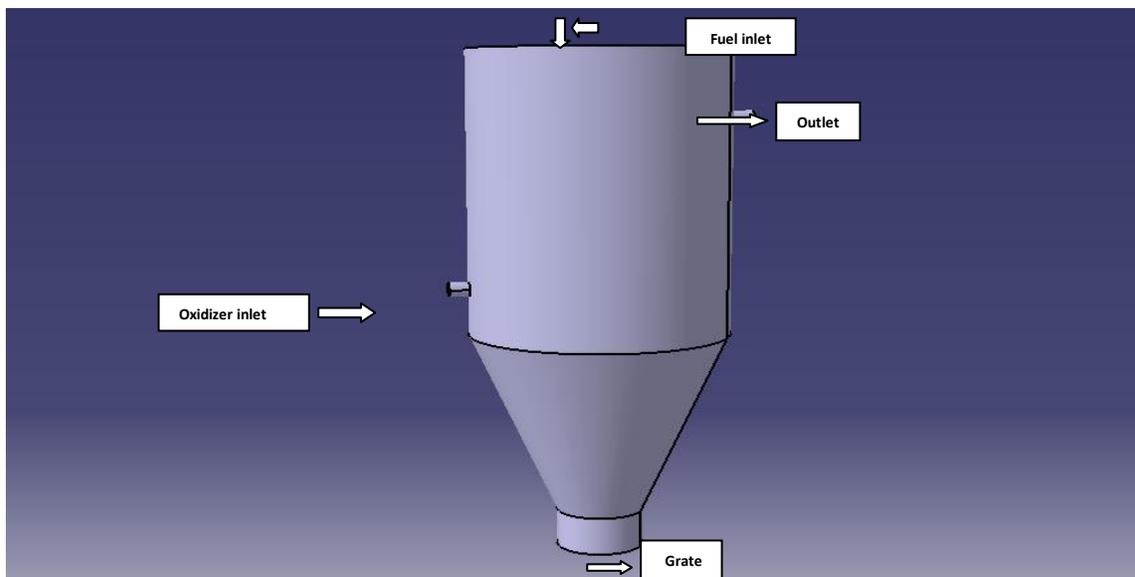


Fig. 1: Model of up-draft gasifier

(ii) Design inputs for simulation of gasifier:

The performance detail of gasifier is given Table 1. Accordingly, with the help of experimental data the performance of gasifier is evaluated in the terms of parameters. The air velocity, fuel rate and moisture content varying different range and more wood

are gasified completely. The ranges of air velocity, fuel rate and moisture content taken 3.7m/s-6.4m/s, 6.9 kg/h-13.6 kg/h and 5%-40% respectively

Table 1: Range of different measured parameters of gasifier.

Measured parameters	Range of parameters
Fuel rate	6.9 kg/h- 13.6 kg/h
Air velocity	3.7-6.4 m/s
Moisture content	5-40 %

(iii)CFD modeling used in this work:

Computation fluid dynamic is a design and analysis tool to simulate fluid flow, heat and mass transfer, chemical reaction, solid and fluid intersection and other related phenomena. Comparing to physical experimental operation, CFD modeling is cost saving, timely, safe and easy to scale – up. The most widely used numerical techniques in CFD are finite difference, finite elements and finite volume methods. Finite volumes are now most commonly used approach in CFD. The most routinely used commercial codes include ANSYS FLUENT, ANSYS CFX, STAR-CD and PHOENICS.

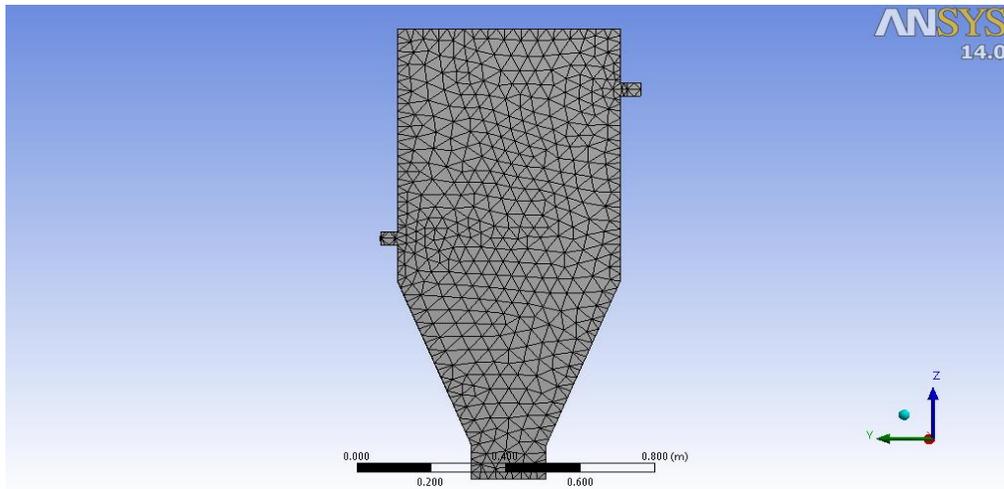


Fig.2: Mesh model of updraft gasifier

The model select of this study focus on the combustion zone of updraft gasifier. CATIA V5R16 software was used to create a three dimensional model of updraft gasifier. ANSYS FLUENT was then used to set the parameters of the model. Finite rate combustion was chosen as the solver for this project. Energy flow model in FLUENT solves the conservation of mass and momentum and the energy equation as well. The species transport solution is solved using the pressure based solver. The Boundary conditions applied for the work is given in the table 2.

Table 2: Boundary condition

Zone	Boundary Type
Oxidizer inlet	Velocity inlet
Output	Pressure outlet
Symmetry	Symmetry
Fuel inlet	Velocity inlet

Velocity inlet: Velocity inlet boundary conditions are used to define velocity and scalar properties of the flow at inlet boundary.

Pressure outlet: Pressure outlet boundary condition is used to define the static pressure at flow outlets. The uses of a pressure outlet boundary condition instead of an outflow condition often result in a better rate of convergence when backflow occurs during iteration.

Mass flow inlet: Mass flow inlet boundary condition are used in compressible flow inlet to prescribe a mass flow rate at the inlet it is not necessary to use mass flow inlets in incompressible flow because when density constant.

IV. Result and Discussions

The simulation was performed on the above model with given boundary conditions. Results are obtained and analyzed. The results analyzed are presented here in three parts. In first part, the effect of air velocity on gas composition, in second part, the effect of fuel rate on gas composition and third part shows the effect of moisture content on gas composition.

(i) Effect of air velocity variation on the value of gas composition

Air velocity is the important variable in gasifier operation. In this analysis air velocity was varied from 3.7 m/s to 6.4 m/s while keeping the fuel feed rate constant at this case. An increased in air velocity suggest an increase oxidation of fuel, and the higher production of CO₂ from element C and more conversion of CO into CO₂. This is also indicated by the decrease mole fraction curve of CO with a simulation increasing in the mole fraction of CO₂. A higher amount CO₂ present in the reactor tends to water gas shift reaction toward greater production of H₂O at higher temperature. When the velocity is increasing 3.7 m/s to 5.4 m/s this is seen in the increasing trend of mole fraction curve of H₂ is 14% to 21% and then they decreased. Table.3 shows the effect of variation of air velocity on gas composition

Table 3: Effect of air velocity on gas composition.

Air Velocity (m/s)	H ₂ (%)	CO (%)	CO ₂ (%)	CH ₄ (%)
3.7	14	25	10	6.3
4.2	15	23	11	6.2
4.6	16	21	12	6.1
5	18	20	14	5.8
5.4	21	19	15	4.9
5.9	16	20	17	4.7
6.4	15	22	19	4.4

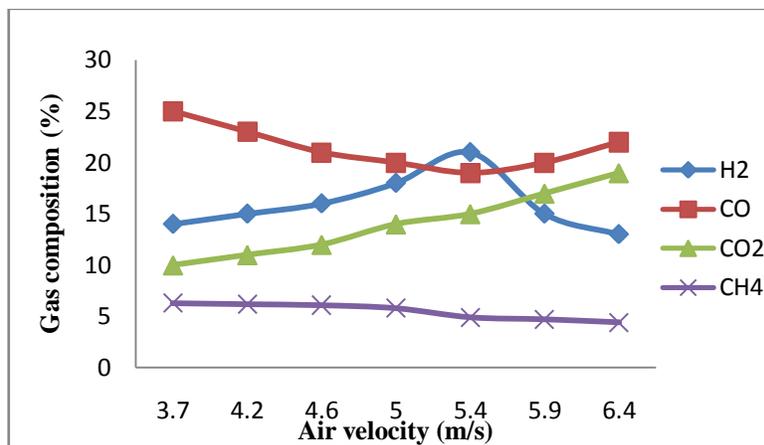


Figure 3: Graph between Air velocity and gas composition

(ii) Effect of fuel rate variation on the value of gas composition:

The input fuel rate to the simulation was varied from 6.9 kg/h to 13.6 kg/h and individual gas composition of producer gas compared with variation of fuel rate. The result from simulation show the figure.4 .when the increased of fuel rate the value of H₂ firstly decreased from 24% to 17% and then increased. And the value of CO first increased from 18% to 22% then decreased by effect of decreasing bed temperature. And CO₂ first decreased then increased by increasing the value of feed fuel velocity.

Table 4: Effect of fuel rate on gas composition.

Fuel rate (kg/h)	H ₂ (%)	CO (%)	CO ₂ (%)	CH ₄ (%)
6.9	24	18	17	4.3
7.4	21	19	16	4.4
8.2	19	21	15	4.9
8.8	17	22	13	6.4
10.7	19	20	14	6.6
11.3	21	18	16	7
13.6	23	16	17	7.1

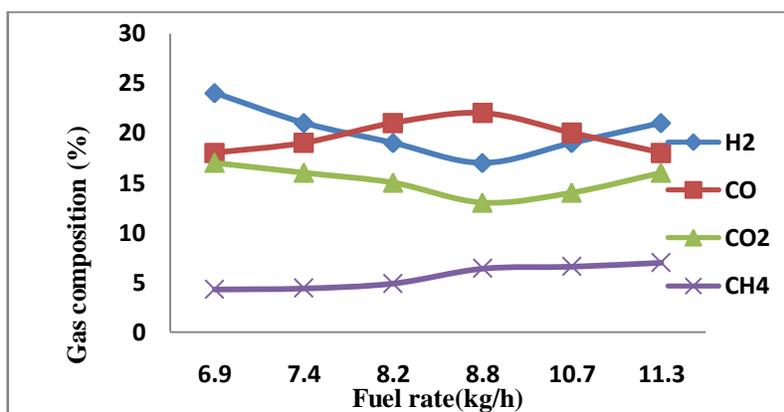


Figure 4. Graph between fuel rate and gas composition

(iii) Effect of moisture content variation on the value of gas composition-

The biomass generally associated with some moisture when feed to the reactor. This analysis was conducted to study the effect of variation of moisture content on the composition of producer gas. This analysis moisture content was varying 5% to 40%. Increasing the level of moisture content of gasification, present CO produce H₂ by water gas shift reaction and this effect increasing H₂ content of producer gas. An increasing CO₂ could also mean increase in combustion due to increased amount of oxidant in the reactor. The concentration of H₂ increased 15% to 24% with change of moisture content from 5% to 35% and then they start to decreased.

Table 5: Effect of moisture content on gas composition

Moisture Content (%)	H ₂ (%)	CO (%)	CO ₂ (%)	CH ₄ (%)
5	15	37	11	4
10	16	35	13	5.5
15	18	34	15	5.6
20	19	31	17	5.7
25	21	31	18	5.8
30	23	34	20	5.8
35	24	39	20	6.7
40	20	41	21	6.8

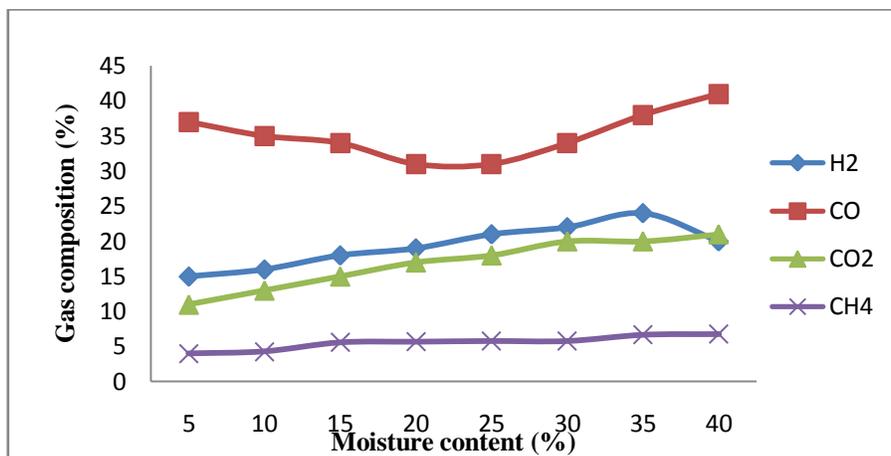


Figure 5. Graph between moisture content and gas composition

V. Conclusions

The detailed CFD model of an up-draft gasifier has been developed, based on FLUENT package. Models of finite rate chemistry in the gas phase and char reactions have been added to the standard model. Simulation performed to predict the gas composition for varying air velocity, fuel rate and moisture content.

- The simulation result shows that, as the *air velocity* increases, the value of H₂ firstly increases from 14% to 21% then decreased and CO concentration first decreased from 25% to 19% then increased. because at one optimum point the air velocity so high that causes on that point gasification not proper done and the value of H₂ decreases and CO increases. CO₂ concentration increased from 10% to 19% because of large amount of element C react with O₂ and produce CO₂. And The Maximum value of H₂ is 21% is obtained at air velocity of 5.4 m/s.
- When increasing *fuel rate*, concentration of H₂ and CO₂ decreased from 24% to 17% and 17% to 13% respectively then increased by increasing the fuel rate and the concentration of CO first increased 18% to 22% then decreased 20% to 16%. Because the air is not enough to gasification that reason the temperature of bed decreased. For decreasing bed temperature value of H₂ start to increase and CO start to decreased. So the optimum value of fuel rate is 8.8 kg/h.
- When increasing the *moisture content* in gasification, the value of H₂ first increased from 15% to 24% then decreased because the level of moisture contents so high. And the value of CO first decreased 37% to 31% and then increased 31% to 41%. And the CO₂ concentration increasing gradually. So the optimum value of moisture content is 25%.

References

- 1) A. Saravanakumar et al. (2007), Experimental investigations of long stick wood gasification in a bottom lit updraft fixed bed gasifier, *Fuel Processing Technology* 88, 617–622.
- 2) A. Saravanakumar et al. (2005), Operation and modeling of an updraft long-stick wood gasifier, *Energy for Sustainable development Volume No.4*.
- 3) Avdhesh Kr. Sharma (2010), Experimental study on 75 kWth downdraft (biomass) gasifier system, *Renewable Energy* 34, 1726–1733.
- 4) Avdhesh kr. Sharma (2010), Modeling and simulation of a downdraft biomass gasifier 1. Model development and validation, *energy conversion and management* 52, (2011) 1386-1396.
- 5) D.F. Fletcher et al. (2000), A CFD based combustion model of an entrained flow biomass gasifier, *Applied Mathematical Modeling* 24, 165-182.
- 6) M. Miltner et al. (2007), Process simulation and CFD calculations for the development of an innovative baled biomass-fired combustion chamber, *Applied Thermal Engineering*, 27, 1138–1143.
- 7) Marta muilenburget al. (2011), Computational modeling of the combustion and gasification zones in a downdraft gasifier. ASME
- 8) P. Abdul Salam et al. (2006), A comparative study of charcoal gasification in two types of spouted bed reactors. *Energy* 31, 228–243
- 9) PengmeiLv et al. (2007), Hydrogen-rich gas production from biomass air and oxygen/steam gasification in a downdraft gasifier, *Renewable Energy* 32, 2173–2185.
- 10) PrabirBasu (2010) Biomass Gasification and Pyrolysis for practical design and theory, *Published by Elsevier*.
- 11) S. Murgia et al. (2011), Comprehensive CFD model of an air-blown coal-fired updraft gasifier, *Fuel*.
- 12) Sirigudirahulrao (2002) Biomass to ethanol: process simulation, validation and sensitivity analysis of a gasifier and a bioreactor, *National Institute of Technology Karnataka*.
- 13) Weihong Yang et al. (2007), Performance analysis of a fixed-bed biomass gasifier using high-temperature air, *Fuel Processing Technology* 87, 235 – 245.
- 14) Yaji Huang et al. (2007), Effect of operating conditions on gas components in the partial coal gasification with air/steam, *Korean J. Chem. Eng.*, 24(4), 698-705.
- 15) Yiqun Wang et al. (2008), Review CFD Studies on Biomass Thermochemical Conversion, *Int. J. Mol. Sci.* 2008, 9, 1108-1130.

Simulation of a Buck-Boost Single Phase Voltage Source Inverter for Distribution Generation Systems

Channaveer. P. Malad¹, K. Uma Rao²

^{1,2}(Department of E&EE, R. V College of Engineering, Bangalore, Karnataka State, India)

Abstract: Inverters of PV system based distributed generators (DG), are often subjected to wide changes in the inverter input voltage, which is either above or below the output ac voltage, thus demanding a buck-boost operation of inverters. Many traditional full-bridge inverters and single-stage buck-boost inverters either have complex structure or have limited range of input dc voltage. In this paper, a single-phase transformer less inverter topology is implemented that can operate over a wide dc input voltage range making it suitable for distributed generation applications. This topology of the inverter has multiple stages and uses six switches. Depending on the reference value set, the inverter output voltage can be either boosted or bucked with respect input voltage. Simulations were carried out using MATLAB simulink software package and results show that the proposed topology boosts or bucks the output voltage level depending upon the value of the reference signal. The necessary capacitor and inductor are easier to design than those for filtering the output of traditional inverters.

Keywords-Buck-Boost inverter, Distribution generators, Full bridge inverter, multiple stages, reference signal.

I. Introduction

Distributed generation (DG) systems are usually small modular devices close to electricity consumers. These include wind turbines, solar energy systems, fuel cells, micro gas turbines, and small hydro systems, as well as relevant control and energy storage systems. Such systems normally need inverters as interfaces between their single phase loads and sources as shown in Figure 1.1, which depicts a renewable energy based DG system[1]-[5]. DG inverters often experience a wide range of input voltage variations due to the fluctuations of energy sources, which impose stringent requirements for inverter topologies and controls.

Functions of inverters for small DG systems can be summarized as follows.

- Power conversion from variable dc voltage into fixed ac voltage for stand-alone applications or ac output in synchronism with the grid voltage and frequency for grid-connected applications. The variable dc voltage can be higher or lower than the ac voltage in a system, which is observed normally in a wind-turbine and solar energy systems. Thus, there is a need to buck or boost the inverter voltage, as the case may be.
- Output of inverter gives power quality assurance with low total harmonic distortion (THD), voltage and frequency deviation, and flickering.
- Protection of DG generators and electric power systems from abnormal voltage, current, frequency and temperature conditions, with

additional functions such as anti islanding protection and electrical isolation if necessary.

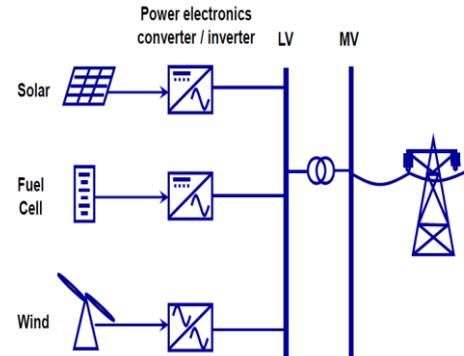


Fig 1: Renewable energy based distribution generation system.

Based on the electrical isolation between the input and output, Inverters can be classified as isolated or non isolated. Electrical isolation is normally achieved using either line-frequency or high-frequency transformers [1].

The dc-link voltage of inverters for DG systems may vary over a wide range. Depending on the input dc voltage range in comparison to the output ac voltage, inverters can be buck inverters, boost inverters, or buck-boost inverters. Different dc voltages are applied to the inverter input because of the new energy sources, such as solar batteries and fuel cells, which produce different dc-voltage levels [4]. Generally, most topologies are boost or buck-boost due to two main factors. First, pulse width modulation (PWM) produces an output voltage lower than the dc link voltage therefore, the dc link should be greater than the maximum possible ac output voltage. Second, sometimes, independent of frequency, voltage step-up is necessary [3].

Traditional full-bridge inverters do not have the flexibility of handling a wide range of dc input voltages. Especially when the dc voltage is lower than the ac voltage, heavy line-frequency step-up transformers are required. Although these inverters demonstrate robust performance and high reliability, they demand higher volume, weight and cost for DG system applications [2], [4].

Buck-boost inverters have the advantage of converting dc voltage higher or lower than the utility voltage without utilizing a line frequency transformer. Two stage or multiple stage configurations are commonly used in buck-boost inverters. Such inverter systems have dc-dc or dc-ac-dc converters added to obtain an elevated dc voltage ahead of inversion. A two-stage buck-boost inverter can achieve a relatively high power capacity; nevertheless, the additional power stage requires more power components and thus higher costs [4].

II. Block Diagram of Buck-Boost Single Phase Voltage Source Inverter For Distribution Generation

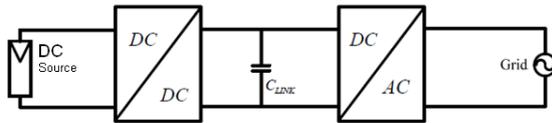


Fig. 2: General Block diagram of buck-boost single phase voltage source inverter for distribution generation

Block diagram of buck-boost inverter single phase voltage source inverter for distribution generation is shown in Fig.2. From block diagram it can be seen that voltage supplied from energy source (Photo voltaic, wind turbines, fuel cell) goes through two stages of conversion before it is supplied to grid. The voltage from the energy source (*dc*) is first converted into variable *dc* using a *dc-dc* converter. This variable *dc* voltage is then converted to *ac* voltage and is given to grid. The capacitor link acts as a voltage source to the inverter

III. Proposed *dc* to *ac* converter

The block diagram of multiple stage buck-boost inverter used for the proposed system is shown in Fig 3. Dc voltage obtained from the photo voltaic cells is given as input to *dc-dc* converter. Depending upon the reference value set, *dc-dc* converter either boosts or bucks the input voltage to 325Vdc. Dc voltage is converted to *ac* voltage by switching the switches of two arms of H-bridge complementarily. The obtained 230V, 50Hz, *ac* voltage is fed to grid.

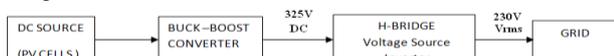


Fig 3: Block diagram of buck-boost voltage source inverter.

The *dc* link voltage is produced by buck-boost converter, prior to being converted to *ac* by the low frequency output H-bridge. A sine reference gives a sinusoidal voltage output that can be used in applications, such as drives, distributed generation, and power systems. Fig 4 gives the schematic of proposed inverter circuit.

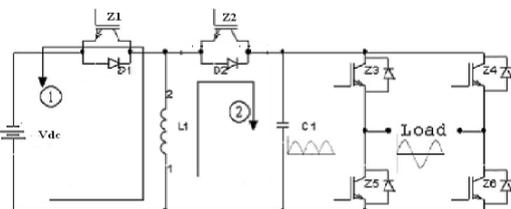


Fig 4: Proposed *dc* to *ac* converter schematic

3.1. Working principle.

For unidirectional operation of the proposed topology of the inverter Z_2 is kept turned OFF throughout the operation. It can be used for bidirectional operation. Switches Z_1 and four switches of H-bridge Z_3, Z_4, Z_5, Z_6 , turned ON and OFF. Average output voltage of buck-boost converter across the capacitor depends upon the duty cycle D . Equivalent circuit for unidirectional buck-boost operation of the converter is shown in Fig 5.

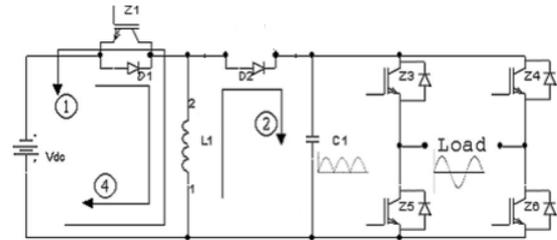


Fig 5: Equivalent circuit for unidirectional buck-boost operation of converter with H-bridge inverter.

When switch Z_1 is ON, input voltage provides energy to inductor and the diode is reversed biased. When switch Z_1 is turned OFF, the energy stored in the inductor is transferred to the capacitor C_1 . The regulated *dc* voltage across the Capacitor C_1 acts as the *dc* link voltage to H-bridge inverter. By switching switches Z_3, Z_6 and Z_4, Z_5 of H-bridge complementarily at power frequency *ac* voltage can be obtained. This voltage can be supplied to grid.

3.2 Flow chart and design consideration for the simulation.

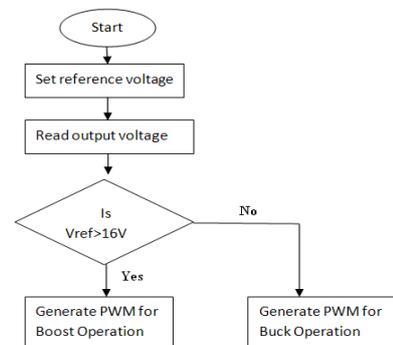


Fig 6: Flow chart to generate PWM for switches Z_1

Flow chart to generate PWM for the switch Z_1 is shown in Fig 6. From flow chart it can be seen that the nominal value of reference sine wave amplitude is 16V. If the reference sine wave amplitude is greater than 16V, the proposed inverter topology acts as a boost inverter whereas if it is less than 16V it acts as a buck inverter. However for if reference sine wave amplitude is equal to 16V, then the proposed topology of the inverter neither boosts the input voltage nor bucks the input voltage, it converts the *dc* input voltage to corresponding output *ac* voltage at 50Hz.

Since the inductor current is discontinuous, the maximum Period for L_1 to deliver energy from or to the *dc* source is limited to half the switching period. Thus, the maximum inductor current $i_{L1 \max}$, which is chosen according to switch and diode maximum current ratings, is given by

$$|i_{L1 \max}| \leq \frac{V_{DC}}{2L_1 f_{sw}} \quad (1)$$

Therefore, the maximum allowable output current i_{omax} (which is different to the maximum rated output current) must be less than. $i_{L1 \max}$

$$|i_{omax}| < |i_{L1 \max}| \quad (2)$$

When L_1 is magnetized from the *dc* source with $i_{L1 \max}$ in the first half of the switching period, L_1 should deliver all its stored energy in the remainder of the period.

The longest time needed for L_1 energy to be released is when C_1 voltage and the output current are both zero. Assuming C_1 voltage is zero and L_1 initial current (the instant the inductor starts to charge the capacitor) is $i_{L1 \max}$, then

$$i_{L1} = i_{L1 \max} \cos\left(\frac{t}{\sqrt{L_1 C_1}}\right) \tag{3}$$

To allow i_{L1} to reach zero, the second period must be greater than or equal to a quarter of the reciprocal of the L_1 - C_1 natural frequency, i.e.

$$f_{sw} \leq \frac{1}{\pi \sqrt{L_1 C_1}} \tag{4}$$

When the inductor transforms energy from the capacitor, back to the dc source, the peak inductor current at the end of the first period should not exceed $i_{L1 \max}$. The worst case occurs when the capacitor voltage is a maximum, then the inductor current is

$$i_{L1} = v_{o\max} \sqrt{\frac{C_1}{L_1}} \sin\left(\frac{t}{\sqrt{L_1 C_1}}\right), \quad \text{for } t \leq \pi \sqrt{L_1 C_1} \tag{5}$$

Thus

$$|i_{L1 \max}| \geq |v_{o\max}| \sqrt{\frac{C_1}{L_1}} \tag{6}$$

IV. Simulations and practical results

The block diagram used for the simulation is shown in the Fig 7. It uses PI and P controller to generate the PWM signal for Z_1 and Z_2 switches. The PWM for H-bridge can given through pulse generator since it is operated at 50Hz.

Simulations have been carried out using MATLAB simulink software package. The proposed topology of the inverter has been implemented using circuit parameters given in Table 1. The input voltage is varied from 100 to 400V (dc) with corresponding change in the amplitude of the reference of sine wave. The circuit working is examined both in boost and buck mode. Simulation circuit for unidirectional buck-boost single phase voltage source inverter is shown in Fig 8. Fig 9 shows the simulink model of H-bridge with LC filter.

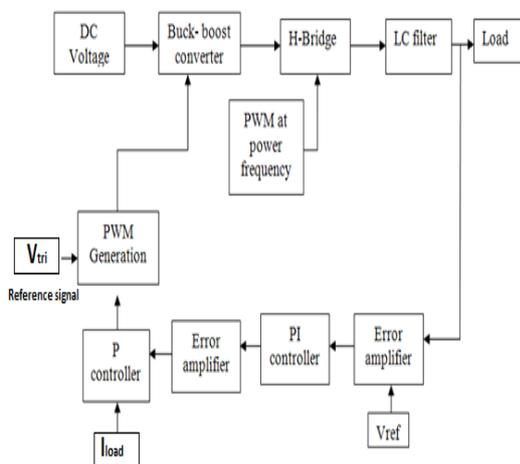


Fig 7: Block diagram of the buck-boost single phase voltage source inverter.

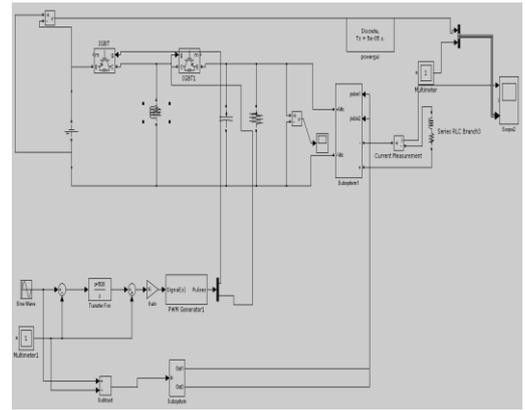


Fig 8: Simulation circuit for unidirectional buck-boost single phase voltage source inverter.

Table 1: System parameters

L_1	100 μ H	P_v	0.3V/A	F_{sw}	20kHz	P_i	0.04
C_1	50 μ F	K_v	600V/A/s	F_s	50Hz		

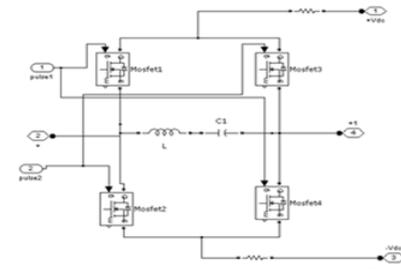


Fig 9: Simulink model of H bridge inverter with LC filter

The proposed topology of the inverter is implemented using circuit parameters given in table 1. The input voltage is varied from 100 to 400V (DC) with corresponding change in amplitude of the reference value of sine wave. The circuit working is examined both in boost and buck mode. Following are the analysis of the same.

4.1 Boost operation.

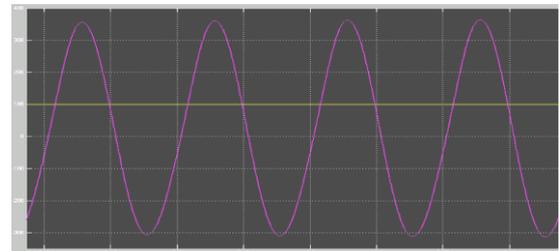


Fig 10: The snap shot of output voltage 650V peak-peak, 50Hz, ac with input voltage of 100V DC and reference sine wave voltage amplitude set at 45V

In this case the proposed topology of the inverter circuit acts a boost inverter since the reference sine wave voltage amplitude applied is 45V. The reference sine wave voltage amplitude set in controller circuit is greater than 16V, the proposed topology the inverter acts as a boost inverter, converting 100V dc input voltage to output voltage of 650V peak-peak ac , 50Hz Fig 10 shows the boost operation of the proposed inverter topology. When input dc voltage is 100V and reference sine wave voltage amplitude set at 45V, the output voltage obtained is 650V peak-peak, 50Hz, ac .

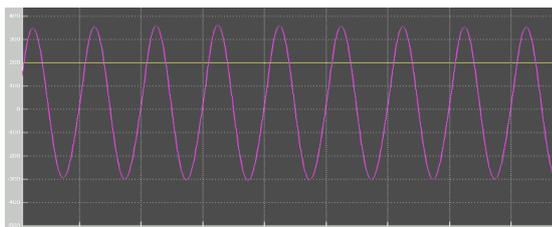


Fig 11: The snap shot of output voltage 650V peak-peak, 50Hz, *ac* with input voltage of 200V DC and reference sine wave voltage amplitude set at 26V.

Fig 11 shows operation of the proposed topology of the inverter circuit, as a boost inverter since the reference sine wave voltage amplitude applied is 26V. The reference sine wave voltage amplitude set in controller circuit is greater than 16V, the proposed topology the inverter acts as a boost inverter, converting 200V *dc* input voltage to output voltage of 650V peak –peak, 50Hz *ac*.

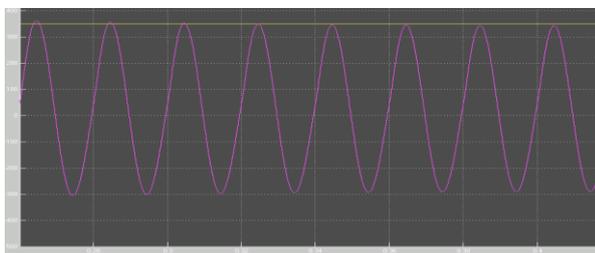


Fig 12: The snap shot of output voltage 650V, peak-peak, 50Hz,*ac*, when input voltage of 325V *dc* is applied and reference sine wave voltage amplitude set at 16V.

Fig 12 shows operation of the proposed inverter circuit topology acting neither as a boost inverter nor a buck inverter since the reference sine wave voltage amplitude applied is 16V. The reference sine wave voltage amplitude given to controller circuit is equal to 16V, which is the nominal value of the voltage for the designed control circuit. For this value of the reference voltage, the proposed topology of the inverter converts input *dc* voltage to corresponding value of output *ac* voltage without boosting or bucking the magnitude input voltage. Thus converting input voltage of 325V *dc* to output voltage of 650V, peak-peak, and 50Hz *ac*.

4.2 Buck operation

Whenever input voltage from PV cells to inverter circuit is greater than grid voltage then the inverter operates in buck mode. In order to operate the inverter in buck mode, amplitude of reference voltage should be set to value less than 16V.

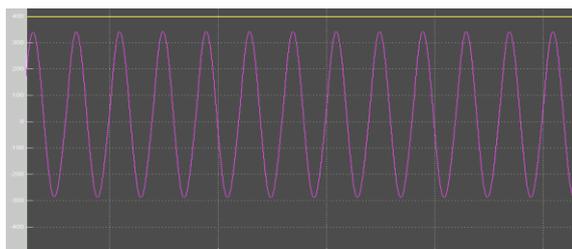


Fig 13: The snap shot of output voltage 650V, peak-peak, 50Hz *ac*, when input voltage of 400V *dc* is applied and reference voltage amplitude set at 13V

Fig 13 shows operation of the proposed topology of the inverter circuit, as a buck inverter since the reference sine wave voltage amplitude applied is 13V. The reference sine wave voltage amplitude set in controller circuit is lesser than 16V, the proposed topology the inverter acts as a buck inverter, converting 400V *dc* input voltage to output voltage of 650V peak –peak, 50Hz, *ac*.

Different output voltages for different input *dc* voltages that are carried during simulation are briefly summarized in Table 2.

Table 2: Summary of simulation results

Sl.No	V _{ref}	V _{dc}	V _{peak-peak}
1	45V	100V	650V
2	32V	150V	650V
3	26V	200V	650V
4	18V	300V	650V
6	16V	325V	650V
7	15V	350V	650V
8	13V	400V	650V

V. Conclusion

Simulation of unidirectional of buck- boost single phase voltage source inverter has been carried out using MATLAB simulink software package. From simulation results it is seen the proposed topology of buck-boost single phase voltage inverter works exceptionally well producing an *ac* sine wave output depending upon the reference sine wave amplitude given to control circuit. The input voltage to power circuit was varied from 100V to 400V. From simulations results it can be seen that if the reference sine wave amplitude is set 16V, which is the nominal value for the simulation control circuit, the voltage source buck-boost inverter neither bucks or boosts the input voltage however it inverts the *dc* voltage to 650V peak-peak 50Hz, *ac* voltage. If the reference sine wave amplitude given to control circuit is greater than 16V then the proposed topology of inverter acts as boost inverter, inverting the input voltage to 650V peak-peak *ac*, 50Hz and vice versa. From simulation results it can be summarized that the proposed topology of the inverter circuit operates for wide range of *dc* input voltage producing a sinusoidal *ac* voltage, 50Hz output.

The proposed inverter is applicable as a utility interactive inverter for distributed generating systems and harmonic elimination applications. The proposed inverter uses six switches. The low switching frequency of the output H-bridge reduces inverter switching losses and costs, compared to six and eight switch-based techniques.

The drawbacks of this inverter, compared to the traditional H-bridge inverter are: relatively high cost (six switches) and relatively high switching losses in two of the six switches.

Reference

- [1] T. G.Wilson, "The evolution of power electronics," *IEEE Trans. Power Electron.*, vol. 15, pp. 439–446, May 2000.
- [2] X. Yaosuo, C. Liuchen, B. K. Kjaer, J. Bordonau, and T. Shimizu, "Topologies of single-phase inverters for small distributed power generators: An overview," *IEEE Trans. Power Electron.*, vol. 19, no. 5, pp. 1305–1314, Sep. 2004.
- [3] R. Gonzalez, E. Gubia, J. Lopez, and L.Maroyo, "Transformerless single phase multilevel-based photovoltaic inverter," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2694–2702, Jul. 2008.

- [4] Ahmed Mohamed Salamah, Stephen J. Finney “Single-Phase Voltage Source Inverter with a Bidirectional Buck–Boost Stage for Harmonic Injection and Distributed Generation” *IEEE transactions on power electronics*, vol. 24, no. 2, pp 376- 387, February 2009.
- [5] Roman, I.T. “A single-phase current-source inverter with active power filter for grid-tied PV systems”, *3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), Aalborg*, pp. 349-356, June 2012.
- [6] T. Boutot and L. Chang, “Development of a single-phase inverter for small wind turbines,” in *Proc. IEEE Elect. Comput. Eng. Can. Conf. (CCECE 1998), Waterloo, ON, Canada*, May 24–28, pp. 305–308.
- [7] P. Lezana, J. Rodriguez, and D. A. Oyarzun, “Cascaded multilevel inverter with regeneration capability and reduced number of switches,” *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 697–703, Mar. 2008.
- [8] Yan Zhang “Comparison of traditional two-stage buck-boost voltage source inverter and diode-assisted buck-boost voltage source inverter” *Twenty-Seventh Annual IEEE trans on Applied Power Electronics Conference and Exposition (APEC), Orlando, FL*, pp 141-148, Feb 2012.
- [9] Samuel, P. “Grid Interface of Wind Power With Large Split-Winding Alternator Using Cascaded Multilevel Inverter” *IEEE Transactions on Energy Conversion*, Volume: 26, Issue: 1, pp 299-309, March 2011.
- [10] Jalbrzykowski, S, Citko, T. “Current-Fed Resonant Full-Bridge Boost DC/AC/DC Converter” *IEEE Transactions on Industrial Electronics*, Volume: 55, Issue:3, pp 1198 – 1205, March 2008.

A Novel Schema for Detecting Malicious Packet Losses

M. Kiran kumar¹, A. Sai Harish²

*(M.Tech, kiet engineering college, Andhra Pradesh, India)

** (Assistant professor, kiet engineering college, Andhra Pradesh, India)

ABSTRACT: Detecting malicious dropping packets is a crucial issue in networks to minimize various security attacks such as blackhole, greyhole, and wormhole attacks. All networks drop the packets in the presence of collisions, channel errors and the network traffic exceeds its capacities. All the existing detection algorithms have addressed this issue by using user-defined threshold value. But these attacks could not able to solve it because too many dropped packets imply malicious intent. To address this problem in this paper we proposed a model to monitor a node for detecting malicious packet dropping. To evaluate the performance of the proposed algorithm we used NS2 simulator. Our experimental reveals that the proposed algorithm performed very well to detect malicious packet drops due to the collisions, channel errors and heavy traffic.

Keywords: channel errors, energy drain attack, malicious dropping attack, collisions, blackhole

I. INTRODUCTION

The Internet environment is not a safe place. Due to the unsecured nodes in the internet even well protected nodes may be face denial-of-service attacks, blackhole, greyhole, and wormhole attacks [1]. However, such attacks to a node are widely understood, it is less well appreciated that the network infrastructure itself is subject to constant attack as well. In this paper, we propose a method to find out where the loss occurred. If a hacker gains the control of a router, he may disturb the communication by dropping or manipulating the transferred packets. Network load can be disturbed by routers, refusing to serve their advertised routes, announcing nonexistent routes, or simply failing to withdraw failed routes, as a result of either malfunction or malice which is described in fig 1. The main idea of detecting malicious packet loss is finding where the packet loss has occurred in the network due to the presence of collisions, channel errors or heavy traffic. The attacker may disturb the packet forwarding by dropping packets routed to it by its neighbors. Mike Lynn's demonstrated how Cisco routers can be compromised via simple software vulnerabilities. Once a router has been compromised in such a fashion, an attacker may interpose on the traffic stream and manipulate it maliciously to attack others by selectively roping, modifying, or rerouting packets.

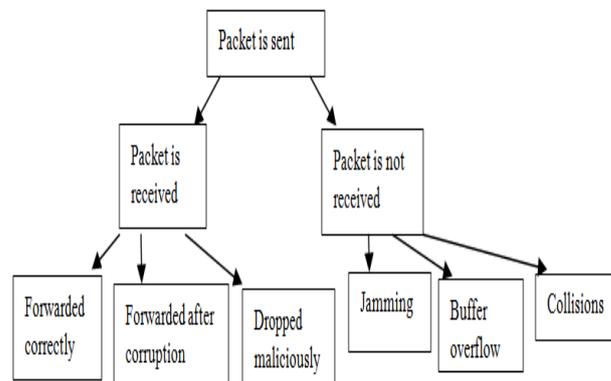


Fig 1: overview of packet loss

Several authors have proposed various protocols to detect packet manipulations, based on validating the traffic transmitted by one router is received unmodified by another [2], [3], [4]. All the proposed algorithms struggle in interpreting the absence of traffic. While a packet that has been modified in transit represents clear evidence of tampering, a missing packet is inherently ambiguous: it may have been explicitly blocked by a compromised router or it may have been dropped benignly due to network congestion. The modern routers drop the packets due to high network traffic and the widely used Transmission Control Protocol (TCP) is designed to cause such losses as part of its normal congestion control behavior [8]. All the existing traffic validation systems must inevitably produce false positives for benign events and/or produce false negatives by failing to report real malicious packet dropping [5]. To overcome this problem, in this paper we proposed a router detection protocol it dynamically infers the precise number of congestive packet losses. If the congestion is avoided, subsequent packet losses can be safely attributed to malicious actions. Our proposed protocol automatically predicts congestion in a systematic manner and takes necessary actions to avoid it. We evaluated the performance of the protocol using NS2 simulator and experimental results relived that the proposed protocol capable of accurately resolving extremely small and fine-grained attacks.

The rest of the paper is organized as section 2: discuss about the related work, section 3: presents the network model, section 4: discuss about the performance metrics, section 5: discuss about the proposed algorithm, section 6: discuss about the experimental setup and section 7: concludes the paper.

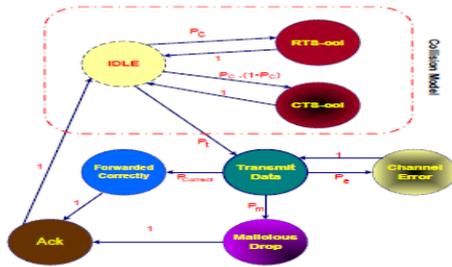


Fig 2: state diagram of packet loss

II. RELATED WORK

H. Ma [9] classified the types of interference which impacts the packet loss in networks. In Type-1 interference, the interference signal arrives before the desired signal. While in Type-2 interference, the interference signal arrives after the desired signal, and in the case of collisions both signals arrive at the same time. Statistical methods have been by used various authors to determine the packet loss rate at each node based on interference type. Pang [2] presented a method to distinguish between packet loss due to collisions and link errors. The main idea is that shorter RTS/CTS and MAC headers in 802.11 are less vulnerable to errors than data. Thus, during the RTS/CTS access procedure, errors are assumed to be due to collisions. If the node receives the CTS frame but not ACK frame then the transmission has more likely failed due to a channel error. However, if an RTS/CTS frame is not received, then the transmission more likely failed due to a collision. If a packet with a corrupted header is received then the receiver will not send anything and the sender assumes a collision as a timeout is occurred. If the data is corrupted, the receiver sends a NAK frame to the sender. But the sender assumes the packet has lost due to channel errors. J. Kim [3] proposed collision aware rate adaptation scheme based on RTS probing to differentiate collisions from channel errors. Malone [5] presented an algorithm to estimate packet losses caused by collisions and by channel errors. This algorithm needs some statics knowledge such as the number of successful transmissions out of the total transmissions over some period of time and the number of slots in which the station does not transmit. S. Marti [10] proposed a watchdog scheme for detecting malicious packet dropping attacks to distinguish between types of packet losses. M. Just [6] has used probes disguised as normal packets to detect malicious nodes and F. Anjum [7] used a centralized authority that receives reports on statistics of various IP flows. But these techniques could not be able to distinguish between causes for packet loss. Appenzeller [25] has explored the question of "How much buffering do routers need?" A widely applied rule-of-thumb suggests that routers must be able to buffer a full delay bandwidth product. Due to congestion control effects, the buffering is proportional to the square root of the total number of TCP flows. To achieve this, the author presented a model of buffer occupancy as a function of TCP behavior.

III. NETWORK MODEL

A state diagram which is shown in fig 2, in the idle state node will be waiting for a packet to send. When a packet arrives, if the medium is free then the node sends an RTS packet. The system may move to an RTS-collision state when two or more nodes that are within each other's range transmit an RTS at the same time with probability P_c . The system may move to a CTS-collision state when a hidden node transmits something that collides with the CTS sent by the receiving node and the CTS-collision will occur with probability $(1 - P_c)P_c$. A node will transmit a packet only if it receives a CTS reply to its RTS. The probability that a data packet is transmitted P_t , is:

$$P_t = (1 - P_c)(1 - P_c(1 - P_c)) = 1 - 2P_c + 2P_c^2 - P_c^3 \quad (1)$$

The size of the RTS/CTS packets is small. After a packet is transmitted it will be either forwarded with probability $P_{correct}$, lost due to channel errors with probability P_e , or maliciously dropped or not ACK-ed with probability P_m . Packets that are maliciously dropped may or may not be acknowledged. In our proposed model we assume that dropped packets will be acknowledged. Node in the Malicious Drop state will move to the ACK state if an ACK message is sent or return to the Transmit state if no ACK is sent. Packets lost due to channel errors will also not be acknowledged by the receiver, and will be retransmitted after some timeout. That is, when the Channel Error state is reached, the node will return back to the Transmit state. If a packet is transmitted by the sender, not dropped due to any errors then the packet will be received or forwarded correctly. Thus the probability to forward a packet correctly $P_{correct}$ can be computed as:

$$P_{correct} = P_t(1 - P_e)(1 - P_m) \quad (2)$$

IV. PERFORMANCE METRICS

Collisions: various authors analyzed the nature of Collisions in 802.11. Bianchi [11] and H. Wu [12] have used Markov chain model to find number of collisions in network. In [13], linearization proposed a method to find an approximate value for P_c based on contention window W and the number of nodes n and is given by:

$$P_c = \frac{2W(n-1)}{(W+1)^2 + 2W(n-1)} \quad (3)$$

Later X. Wang [14] presented a approach based on probability of a node sends the packet. $P_c = 1 - (1-T)^{n-1}$ where n is the average number of contending nodes and T is the average probability that a node sends a packet T is denoted as $1/w$ and the probability of collision can defined as :

$$P_c = 1 - (1 - 1/W)^{n-1} \quad (4)$$

Finally Using the number of RTS and CTS packets that were counted during a time window w the probability that a packet was lost due to collision can defined as :

$$P_c = (\#RTS - \#CTS) / \#RTS \quad (5)$$

Channel errors: TN. Gupta [15] assumed a wireless channel with Markov chain model to analyze the performance of 802.11. The duration of wireless channel good and bad state is defined as λ_g^{-1} and λ_b^{-1} , respectively. J. N. ArauzIn [16] performed several experiments and modeled 802.11 links to find the values of λ_g^{-1} and λ_b^{-1} for several PHY layer bit rates and three SNR levels (high, medium, and low).

Energy Drain Attack: In this attack, malicious node intends to drain the sender's battery by not sending ACKs and making the sender retransmit the packet several times before sending an ACK. When the malicious node responds with an ACK to a data packet, the sender node will assume that the packet has been received and forwarded correctly. In this case, the sender node estimates $P'_{Correct}$ as:

$$P'_{Correct} = \#ACK/\#RTS \quad (6)$$

Malicious node may drop the ACK-ed packet and not relay it to the next hop, because the attack is directed towards draining the battery, the ultimate fate of the ACK-ed packet is not relevant.

Malicious Dropping Attack: In this attack, malicious receiving node may send an ACK message upon receiving a packet to be relayed and not forward the packet to the next hop. There are two possible ways to know if the acknowledged packet was forwarded or not, either by monitoring the node using overhearing capability or by having feedback from intermediate nodes which include communication overhead. Hence, in the proposed model we prefer to monitor the receiving node.

V. PROPOSED ALGORITHM

The proposed algorithm used to detect if the neighbor node maliciously dropping packets.

Step 1: Node A will count the RTS messages it sent to node B during some time window w and also the CTS messages received from node B during the same time.

Step 2: Node A will use the model previously described for the value of P_e based on the link SNR. We assume symmetric links, and thus the SNR is expected to be equal at the sending and receiving sides.

Step 3: If the goal of node A is only to prevent energy drain attacks then compute $P'_{Correct}$.

Step 4: If the goal of node A to detect malicious packet dropping then it will use monitoring through overhearing to get an estimate of $P'_{Correct}$.

Step 5: Node A calculate the percentage of packets being maliciously dropped. If P_m is greater than some threshold value then the node is marked as being malicious and node A will inform other neighbors, remove it from routes, etc..

VI. EXPERIMENTAL SETUP

We have implemented our proposed algorithm in NS2, which has been highly validated by the networking research community. The simulation parameters were listed in table 1.

Table 1: NS2 parameters

Parameters	Value
MAC Layer	IEEE 802.11
Number of nodes	20
Data rate	11Mbps
Packet Size	512 B
Simulation Duration	200 sec
Traffic Flow	TCP

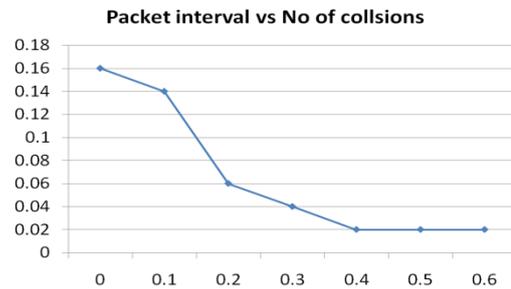


Fig 3: packet interval vs no of collisions

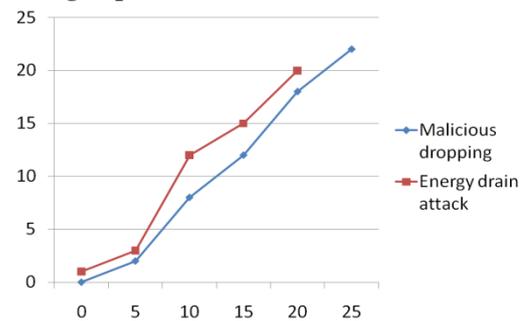


Fig 4: shows the probability of collisions for different traffic loads.

Fig 3 describes the less packet interval time, the more is the load so high probability of collision. Fig 4 shows the "computed" P_m percentage values at each node for the energy drain attack (ED) and malicious dropping (MD) as a function of simulation "specified" malicious packet (or ACK) dropping levels.

VII. CONCLUSION

Detecting malicious dropping packets is a crucial issue in networks to minimize various security attacks such as blackhole, greyhole, and wormhole attacks. All networks drop the packets in the presence of collisions, channel errors and the network traffic exceeds its capacities, which depend on the environment of the network. Hence, in this paper we present a method to determine the cause of packet drops by a node such as collisions, channel errors and heavy traffic conditions. If nodes can have reasonable estimates for collision probabilities and channel error probabilities, even fairly low levels of malicious packet drops can be detected significantly. To evaluate the performance of the proposed method, we simulated it using NS2. Experimental results relived that proposed method performances well.

REFERENCES

- [1] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Ariadne: a secure on-demand routing protocol for ad hoc networks," *Wirel. Netw.*, vol. 11, no. 1-2, pp. 21–38, 2005.
- [2] Q. Pang, S. Liew, and V. Leung, "Design of an effective lossdistinguishable mac protocol for 802.11 wlan," *Communications Letters, IEEE*, vol. 9, no. 9, pp. 781–783, Sep 2005.
- [3] J. Kim, S. Kim, S. Choi, and D. Qiao, "Cara: Collision-aware rate adaptation for ieee 802.11 wlangs," *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pp. 1–11, April 2006.
- [4] J.-H. Yun and S.-W. Seo, "Novel collision detection scheme and its applications for ieee 802.11 wireless lans," *Computer Communications*, vol. 30, no. 6, pp. 1350–1366, —2007.
- [5] D. Malone, P. Clifford, and D. J. Leith, "Mac layer channel quality measurement in 802.11," *Communications Letters, IEEE*, vol. 11, no. 2, pp. 143–145, Feb. 2007.
- [6] M. Just, E. Kranakis, and T. Wan, "Resisting malicious packet dropping in wireless ad hoc networks," in *In Proc. of ADHOCNOW03*. Springer Verlag, 2003, pp. 151–163.
- [7] F. Anjum and R. Talpade, "Lipad: lightweight packet drop detection for ad hoc networks," *Vehicular Technology Conference, 2004. VTC2004- Fall. 2004 IEEE 60th*, vol. 2, pp. 1233–1237 Vol. 2, Sept. 2004.
- [8] O. F. Gonzalez, M. P. Howarth, and G. Pavlou, "Detection of packet forwarding misbehavior in mobile ad-hoc networks." in *WWIC,ser. Lecture Notes in Computer Science*, F. Boavida, E. Monteiro,
- [9] S. Mascolo, and Y. Koucheryavy, Eds., vol. 4517. Springer, 2007, pp. 302–314. [Online]. Available: <http://dblp.uni-trier.de/db/conf/wwic/wwic2007.html#GonzalezHP07>.
- [10] H. Ma, J. Zhu, and S. Roy, "On loss differentiation for csma-based dense wireless network," *Communications Letters, IEEE*, vol. 11, no. 11, pp. 877–879, November 2007.
- [11] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2000, pp. 255–265.
- [12] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 3, pp. 535–547, Mar 2000.
- [13] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over ieee 802.11 wireless lan: analysis and enhancement," *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 599–607 vol.2, 2002.
- [14] M. Carvalho and J. Garcia-Luna-Aceves, "Delay analysis of ieee 802.11 in single-hop networks," *Network Protocols, 2003. Proceedings. 11th IEEE International Conference on*, pp. 146–155, Nov. 2003.
- [15] J. Yin, X. Wang, and D. Agrawal, "Optimal packet size in error-prone channel for ieee 802.11 distributed coordination function," *Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE*, vol. 3, pp. 1654–1659 Vol.3, March 2004.
- [16] N. Gupta and P. R. Kumar, "A performance analysis of the 802.11 wireless lan medium access control," *Communications in Information and Systems*, vol. 3, no. 4, pp. 279–304, 2004.
- [17] J. N. Arauz, "802.11 markov channel modeling," Ph.D. dissertation, School of Information Sciences, University of Pittsburgh, 2004.
- [18] Thaier Hayajneh, Prashant Krishnamurthy, David Tipper, and Taehoon Kim, "Detecting Malicious Packet Dropping in the Presence of Collisions and Channel Errors in Wireless Ad hoc Networks," *IEEE Transaction*, 2009.

Power Quality Improvement in Wind Energy system by using STATCOM on Integration to the Grid

V. Amarnath Reddy¹, P. Harshavardhan Reddy², M. Sudheer babu³

^{1,3} M. Tech Studet, Department of EEE, ²Assistant Professor, Department of EEE
A.I.T.S Engineering College, Rajampeta, India.

Abstract: Renewable energy sources are alternative energy source, can bring new challenges when it is connected to the power grid. Generated power from wind energy system is always fluctuating due to the fluctuations in the wind. According to the guidelines specified in IEC-61400 standard (International Electro-technical Commission) provides some norms and measurements. The performance of the wind turbine, power quality is determined. The power quality measurements are-the active power, reactive power, voltage sag, voltage swell, flicker, harmonics, and electrical behavior of switching operation and these are measured according to national/international guidelines. The paper clearly shows the existence of power quality problem due to installation of wind turbine with the grid. In this STATCOM is used with energy storage system (BESS) to reduce the power quality problems. The STATCOM control scheme for the grid connected wind energy generation system to improve the power quality is simulated using MATLAB/SIMULINK in power system block set.

Keywords: Power Quality, Wind Generating System (WGS), STATCOM, BESS, IEC standard.

I. INTRODUCTION

The need to integrate the renewable energy like wind energy into power system is to minimize the environmental impact on conventional plant. The integration of wind energy into existing power system presents requires the consideration of voltage regulation, stability, power quality problems. The power quality is an essential customer-focused measure and is greatly affected by the operation of a distribution and transmission network.

The individual units can be of large capacity up to 2 MW wind turbine, feeding into distribution network, particularly with customers connected in close proximity. Today, more than 28 000 wind generating turbines are successfully operating all over the world. Fluctuations in the wind speed are transmitted as fluctuations in the mechanical torque, electrical power on the grid and leads to large voltage fluctuations. Fluctuations in network, such as voltage sag, swells, flickers, harmonics etc.

The proposed STATCOM control scheme for grid connected wind energy generation for power quality improvement has following objectives.

- Maintains Unity power factor at source side.
- Supports Reactive power only from STATCOM to wind Generator and Load.
- Simple PI controller for STATCOM to achieve fast dynamic response.

The paper is organized as follows. The Section II introduces the power quality standards, issues and its consequences of wind turbine and the grid coordination rule

for grid quality limits. The Section III describes the topology for power quality improvement. The Sections IV, V, VI describes the control scheme, system performance and conclusion respectively.

II. POWER QUALITY IMPROVEMENT

A. Power quality standards, issues and its consequences

1) International electro technical commission guidelines: Some guidelines of measurements and norms are specified under IEC 61400 standard which determines the power quality of wind turbines.

The standard norms are specified.

- 1) IEC 61400-21: Measuring the power quality characteristic of grid connected wind turbine.
- 2) IEC 61400-13: Wind Turbine—measuring procedure in determining the power behavior.
- 3) IEC 61400-3-7: Measures the emission limits for fluctuating load and IEC 61400-12: Wind Turbine performance.

2) Harmonics: It is due to the operation of power electronic converters. Harmonic voltage and current should be in limited as per the IEC-61400-36 guideline. The rapid switching gives a large reduction in lower order harmonic current and higher order harmonics are filtered out by using filters.

3) VOLTAGE VARIATION: This is due to the fluctuations in the wind turbine due to wind. The voltage variation is directly related to real and reactive power variations. The voltage variation is commonly classified as under:

- Voltage Sag/Voltage Dips.
- Voltage Swells.
- Short Interruptions.
- Long duration voltage variation.

The voltage flicker issue describes dynamic variations in the network caused by wind turbine or by varying loads. Amplitude of voltage fluctuations depends on grid strength, network impedance, phase angle and power factor of wind turbine.

During voltage variations frequency is in the range 10–35 Hz. The IEC 61400-4-15 specifies a flicker meter that can be used to measure flicker directly.

4) WIND TURBINE LOCATION IN POWER SYSTEM: It is located where the power quality is highly influenced.

Its operation and its influence on the power system depend on the structure of the network.

5) SELF EXCITATION OF WIND TURBINE

GENERATING SYSTEM: The self-excitation of wind turbine generating system (WTGS) arises a risk equipped with commutating capacitor. It provides the reactive power compensation to the induction generator.

The disadvantages of self-excitation are the safety aspect and balance between real and reactive power.

6) CONSEQUENCES OF THE ISSUES: Voltage variations, voltage flicker, harmonics causes the malfunctions of equipments. It leads to tripping of protection devices, damaging the sensitive equipments. Overall it degrades the power quality in the grid.

B. GRID COORDINATION RULE

American Wind Energy Association (AWEA) led the effort to develop its own grid code for stable operation as per IEC-61400-21 for the interconnection of wind plants to the utility systems, after the blackout in United State in August 2003.

According to these, operator of transmission grid is responsible for the organization and operation of interconnected system.

1) Voltage rise (u)

The voltage rise at the point of common coupling can be approximated as a function of maximum apparent power S_{max} of the turbine, the grid impedances R and X at the point of common coupling and the phase angle Φ , given in Eq. 1.

$$\Delta u = \frac{s_{max} (R \cos \Phi - X \sin \Phi)}{u^2} \quad (1)$$

Where Δu —voltage rise,
 s_{max} —max. apparent power,
 Φ —phase difference,
 U —nominal voltage of grid.

The Limiting voltage rise value is <2 %

2) Voltage dips (d)

The voltage dips is due to startup of wind turbine and it causes a sudden reduction of voltage. It is the relative % voltage change due to switching operation of wind turbine. The decrease of nominal voltage change is given in Eq. 2.

$$D = K_u \frac{s_n}{s_k} \quad (2)$$

Where d is relative voltage change,
 s_n is rated apparent power,
 s_k is short circuit apparent power, and
 K_u is sudden voltage reduction factor.
 The acceptable voltage dips limiting value is <3%.

3) Flicker

The measurements are made for maximum number of specified switching operation of wind turbine with 10-min period and 2-h period are specified, as given in Eq. 3.

$$P_u = c (\psi_k) \frac{s_n}{s_k} \quad (3)$$

Where P_u —Long term flicker.
 $c (\psi_k)$ —Flicker coefficient

The Limiting Value for flicker coefficient is about ≤ 0.4 , for average time of 2 h.

4) Harmonics

The harmonic distortion is assessed for variable speed turbine with a electronic power converter at the point of common connection. The total harmonic voltage distortion of voltage is given as in Eq. 4.

$$V_{THD} = \sqrt{\sum_{h=2}^{40} \frac{V_n^2}{V_1^2}} \cdot 100 \quad (4)$$

Where V_n is the nth harmonic voltage and V_1 is the fundamental frequency (50) Hz.

The THD limit for 132 KV is < 3%.

THD of current I_{THD} is given as in Eq. 5

$$I_{THD} = \sqrt{\sum_{h=2}^{40} \frac{I_n^2}{I_1^2}} \cdot 100 \quad (5)$$

where I_n is the nth harmonic current and I_1 is the fundamental frequency (50) Hz.

The THD of current and limit for 132 KV is <2.5%.

5) GRID FREQUENCY

The grid frequency in India is specified in the range of 47.5–51.5 Hz, for wind farm connection.

III. TOPOLOGY FOR POWER QUALITY IMPROVEMENT

The STATCOM based current control voltage source inverter injects the current into the grid will cancel out the reactive part and harmonic part of the load and induction generator current, thus it improves the power factor and the power quality. To accomplish these goals, the grid voltages are sensed and are synchronized in generating the current.

The proposed grid connected system is implemented for power quality improvement at point of common coupling (PCC), for grid connected system in Fig. 1

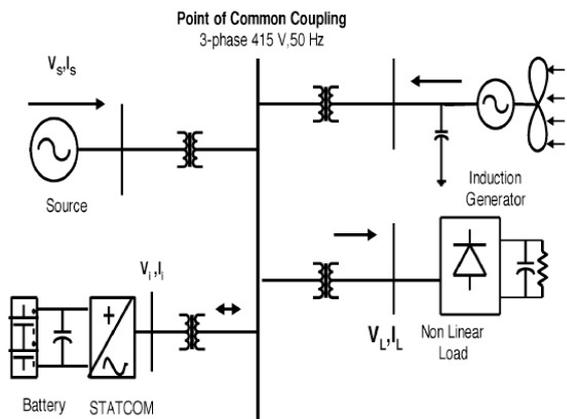


Fig.1. Grid connected system for power quality improvement.

A. WIND ENERGY GENERATING SYSTEM: In this configuration, wind generations are based on constant speed topologies with pitch control turbine.

The induction generator is used in the proposed scheme because of its simplicity, it does not require a separate field circuit, it can accept constant and variable loads, and has natural protection against short circuit.

The available power of wind energy system is presented as under in Eq.6.

$$P_{wind} = \frac{1}{2} \rho A V_{wind}^3 \quad (6)$$

Where ρ (kg/m) is the air density and
 A (m) is the area swept out by turbine blade,
 V wind is the wind speed in mtr/s.

It is not possible to extract all kinetic energy of wind, thus it extract a fraction of power in wind, called power coefficient C_p of the wind turbine, and is given in Eq.7

$$P_{mech} = C_p P_{wind} \quad (7)$$

Where C_p is the power coefficient, depends on type and operating condition of wind turbine. This coefficient can be express as a function of tip speed ratio γ and θ pitch angle. The mechanical power produce by wind turbine is given in Eq. 8.

$$P_{mech} = \frac{1}{2} \rho \pi R^2 V_{wind}^3 C_p \quad (8)$$

Where R is the radius of the blade (m).

B. STATCOM – STATIC SYNCHRONOUS COMPENSATOR

The STATCOM (or SSC) is a shunt-connected reactive-power compensation device that is capable of generating and/ or absorbing reactive power and in which the output can be varied to control the specific parameters of an electric power system.

In general it is solid state switching converter which is capable of generating or absorbing independently controllable real and reactive power at its output terminals when it is fed from an energy source at its input terminals.

Specifically, the STATCOM considered in this is a voltage-source converter from a given input of dc voltage produces a set of 3-phase ac-output voltages, each in phase with and coupled to the corresponding ac system voltage through leakage reactance.

The dc voltage is provided by an energy-storage capacitor.

A STATCOM can improve power-system performance in such areas as the following:

1. The dynamic voltage control in Transmission and distribution systems;
2. The power-oscillation damping in power transmission systems;
3. The transient stability;
4. The voltage flicker control; and
5. It also controls real power in line when it is needed.

Advantages

- 1) It occupies small areas.
- 2) It replaces the large passive banks and circuit elements by compact converters.
- 3) Reduces site work and time.
- 4) Its reponse is very fast.

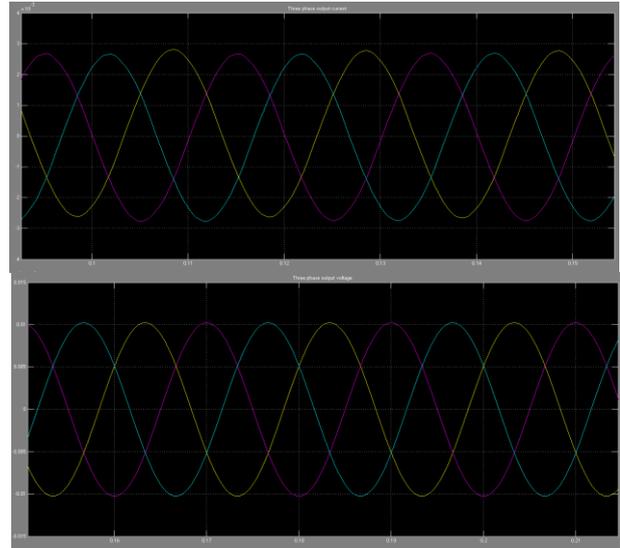


Fig.3. 3phase output current and voltage of grid

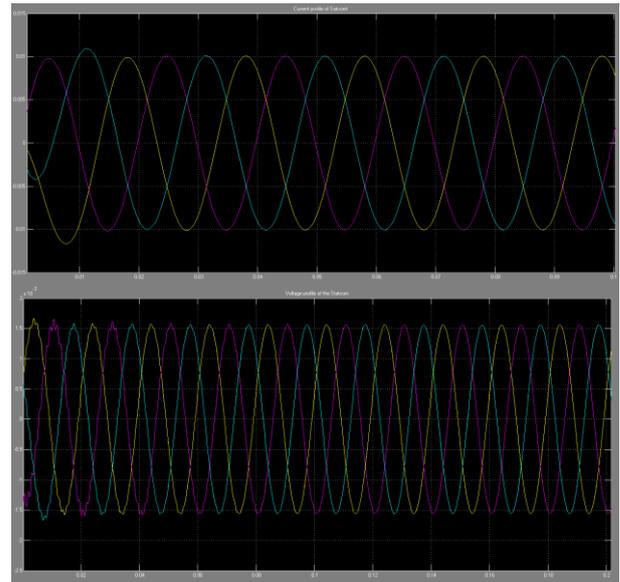


Fig.4. 6-pulse statcom output current and voltage

IV. CONTROLLER DESIGN OF PI, PD AND PID

It is possible to improve the STATCOM response by employing the PID control method by choosing k_p , k_I and k_D . It is a time consuming process but response speed, settling time and proper overshoot rate all guarantees the system stability.

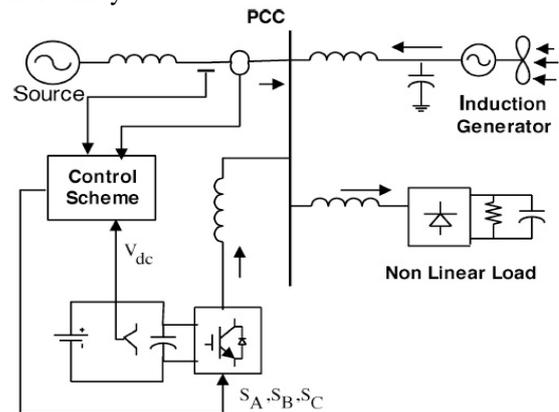


Fig. 2. System operational scheme in grid system.

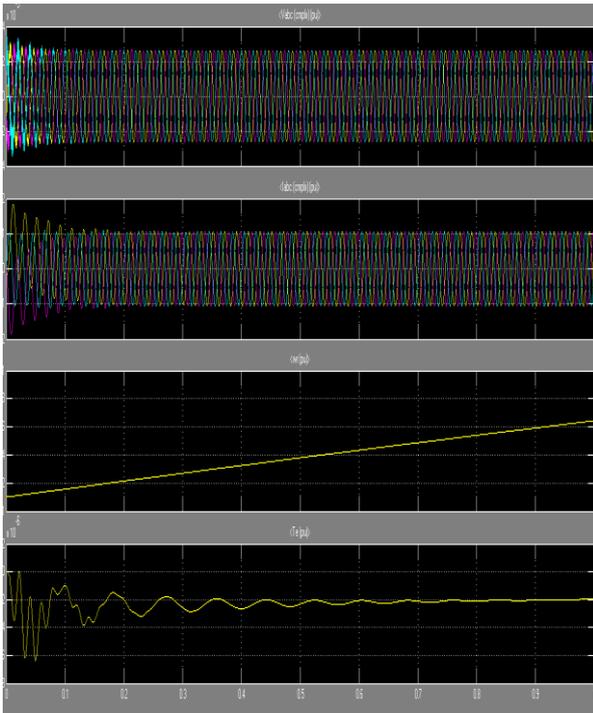


Fig 5. Voltage and current profile of wind generator

V. SYSTEM PERFORMANCE

The proposed control scheme is simulated using SIMULINK in power system block set. The system parameter for given system is given Table I.

	Parameters	Ratings
1	Grid Voltage	3-Phase, 415V,50Hz
2	Induction motor/generator	3.35KVA, 415V,Hz,P=4, Speed=1440rpm,Rr=0.01Ω, Rs=0.015Ω,Ls=Lr=0.06H
3	Line series inductance	0.05mH
4	Inverter Parameters	DC Link Voltage=800V, DC Link capacitance=100μF, Switching Frequency=2kHz
5	IGBT rating	Collector Voltage=1200V, Forward Current=50A,Gate Voltage=20V,Power Dissipation=310w
6	Load Parameter	Non-Linear Load=25kw

VI. CONCLUSION

In this paper we present the FACTS device (STATCOM) -based control scheme for power quality improvement in wind generating system on integration to the grid and with nonlinear load. The power quality issues and its consequences on the consumer and electric utility are presented. The operation of the control system developed for the STATCOM in MATLAB/SIMULINK for maintaining the power quality is to be simulated. It has a capability to cancel out the harmonic parts of the load

current. It maintains the source voltage and current in-phase and support the reactive power demand for the wind generator and load at PCC in the grid system, thus it enhance the utilization factor of transmission line. The integrated wind generation and FACTS device with BESS have shown the outstanding performance.

Thus the proposed scheme in the grid connected system fulfills the power quality norms as per the IEC standard 61400-21.

ACKNOWLEDGMENT

The authors wish to acknowledge the support provided by Annamacharya institute of Technology and Sciences, Rajampet to complete the work successfully.

REFERENCES

1. A STATCOM-Control Scheme for Grid Connected Wind Energy System for Power Quality Improvement Sharad W. Mohod, *Member, IEEE*, and Mohan V. Aware
2. A. Sannino, "Global power systems for sustainable development," in *IEEE General Meeting*, Denver, CO, Jun. 2004.
3. K. S. Hook, Y. Liu, and S. Atcitty, "Mitigation of the wind generation integration related power quality issues by energy storage," *EPQU J.*, vol. XII, no. 2, 2006.
4. R. Billinton and Y. Gao, "Energy conversion system models for adequacy assessment of generating systems incorporating wind energy," *IEEE Trans. on E. Conv.*, vol. 23, no. 1, pp. 163–169, 2008, Multistate.
5. *Wind Turbine Generating System—Part 21*, International standard-IEC61400-21, 2001.
6. J. Manel, "Power electronics system for grid integration of renewable energy source: A survey," *IEEE Trans. Ind. Electron.*, vol.53,no.4,pp 1002-1014, 2006,carrasco.
7. M. Tsili and S. Papathanassiou, "A review of grid code technology requirements for wind turbine," *Proc. IET Renew. Power gen.*, vol. 3, pp. 308–332, 2009.
8. S. Heier, *Grid Integration of Wind Energy Conversions*. Hoboken, NJ: Wiley, 2007.
9. J. J. Gutierrez, J. Ruiz, L. Leturiondo, and A. Lazkano, "Flicker measurement system for wind turbine certification," *IEEE Trans. In strum. Meas.*, vol. 58, no. 2, pp. 375–382, Feb. 2009.
10. Indian Wind Grid Code Draft report on, Jul. 2009, pp. 15–18, C-NET

Supply Chain Production Inventory Model: Innovative Study for Shortages Allowed With Partial Backlogging

Jasvinder Kaur¹, Rajendra Sharma²

*(Department Of Mathematics, Graphic Era University Dehradun, India)

** (Department Of Mathematics Graphic Era University Dehradun, india)

Abstract: In this paper, we have strived to combine all the above mentioned factors into a single problem. We shall undertake to explore a two echelon supply chain, comprising of a vendor and a buyer. The whole environment of business dealings has been assumed to be progressive credit period, which conforms to the practical market situation. The whole combination is very unique and very much practical. The variable holding cost and variable setup has been explored numerically as well; an optimal solution has been reached. The final outcome shows that the model is not only economically feasible, but stable also.

Keywords: Inventory model, partial backlogging, Progressive permissible delay, Supply chain, Shortages, EOQ (Economics Order Quantity).

I. INTRODUCTION

Inventory represents one of the most significant possessions that most businesses possess. It is in direct touch with the user department in its day today activities. Inventory management is playing a key role in setting up efficient closed loop supply chains. A supply chain is a network of facilities and distribution options that performs the functions of procurement of materials, transformation of these materials into intermediate and finished products, and the distribution of these finished products to customers. It consists of a network of companies which are dependent on each other while making independent decisions. The supply chain not only includes the manufacturer and suppliers, but also transporters, warehouses, retailers, and customers themselves. Therefore, supply chain analysis tools and methodologies have become more and more important. It can be a source of great efficiency and cost-savings gains. Supply chain speed and flexibility have become key levers for competitive differentiation and increased profitability. The faster the supply chain, the better a company can respond to changing market situation and the less it needs inventory which resulting in higher return on capital employed. Supply chain management offers a large potential

or organizations to reduce costs and improve customer service performance. In the existing literature, most of the inventory models studies only aimed at the determination of the optimum solutions that minimized cost or maximized profit from the vendor's and vendor's side. However, in the modern global competitive market, the buyer and vendor should be treated as strategic partners in the supply chain with a long term cooperative relationship. Recently, many researchers have considered the buyer and vendor as a unit to find the optimal EOQ in achieving the minimum total cost. In today's business transactions, it is more and more common to see that the customers are allowed some grace

Period before they settle the account with the supplier. This provides an advantage to the customers, due to the fact that they do not have to pay the supplier immediately after receiving the product, but instead, can defer their payment until the end of the allowed period. The customer pays no interest during the fixed period they are supposed to settle the account; but if the payment is delayed beyond that period, interest will be charged. The customer can start to accumulate revenues on the sale or use of the product and earn interest on that revenue. So it is to the advantage of the customer to offer the payment to the supplier until the end of the period.

The two famous formulae of EOQ and EPQ are treated separately for a buyer and a vendor respectively. From the traditional point of view, the vendor and the buyer are two individual entities with different objectives and self-interest. Due to rising costs, the globalization trend, shrinking resources, shortened product life cycle and quicker response time, increasing attention has been placed on the collaboration of the whole supply chain system. An effective supply chain network requires a cooperative relationship between the vendor and the buyer. It assumes that the buyer must pay off as soon as the items are received. Suppliers often offer trade credit as a marketing strategy to increase sales and reduce on-hand stock is reduced, and that leads to a reduction in the buyer's holding cost of finance. In addition, during the time of the credit period, buyers may earn interest on the money. In fact, buyers, especially small businesses which tend to have a limited number of financing opportunities rely on trade credit as a source of short-term funds. The classical inventory models have considered demand rates which were either constant or depended upon a single factor only, like, stock, time etc. But changing market conditions have rendered such a consideration quite unfruitful, since in real life situation, a demand cannot depend exclusively on a single parameter. A combination of two or more factors grants more authenticity to the formulation of the model. Many delivery policies have been proposed in literature for this problem. **Clark and Scarf (1960)** presented the concept of serial multi-echelon structures to determine the optimal policy. **Goyal (1985)** considered a mathematical models with a permissible delay in payments to determine the optimal order quantities. **Ha and Kim (1997)** used a graphical method to analyze the integrated vendor-buyer inventory status to derive an optimal solution. **Hwang and Shinn (1997)** studied effects of permissible delay in payments on retailer's pricing and lot sizing policy for exponentially deteriorating products. **Yang and Wee (2000)** developed an integrated economic ordering policy of deteriorating items for a vendor and a buyer. **Wang et al. (2000)** analyzed supply chain models for perishable products under inflation and permissible delay in payment.

Teng (2002) modified **Goyal (1985)** model by considering the selling price, instead of purchasing cost, as the base to calculate the interest. **Abad and Jaggi (2003)** studied a seller-buyer model with a permissible delay in payments by game theory to determine the optimal unit price and the credit period, considering that the demand rate is a function of retail price. **Huang, Y.F. et al. (2005)** considered the optimal inventory policies under permissible delay in payments depending on the ordering quantity. **Song and Cai (2006)** has been taken on optimal payment time for a retailer under permitted delay of payment by the wholesaler. **Liao (2007)** assumed on an EPQ model for deteriorating items under permissible delay in payments.

In the present study, we have strived to combine all the above mentioned factors into a single problem. We shall undertake to explore a two echelon supply chain, comprising of a vendor and a buyer. The whole environment of business dealings has been assumed to be progressive credit period, which conforms to the practical market situation. The whole combination is very unique and very much practical. The variable holding cost and variable setup has been explored numerically as well; an optimal solution has been reached. The final outcome shows that the model is not only economically feasible, but stable also.

II. PROPOSED ASSUMPTIONS & NOTATIONS

1. ASSUMPTIONS

The following assumptions are used to develop aforesaid model:

- 1.1 The demand rate, $D(t)$, is deterministic, the demand function $D(t)$ is given by $D(t) = \lambda_0 e^{\alpha t}$, a and b are positive constants.
- 1.2 Shortages are allowed with partial backlogging.
- 1.3 If the retailer pays by M , then the supplier does not charge to the retailer. If the retailer pays after M and before N ($N > M$), he can keep the difference in the unit sale price and unit purchase price in an interest bearing account at the rate of I_c /unit/year. During $[M, N]$, the supplier charges the retailer an interest rate of I_{c1} /unit/year on unpaid balance. If the retailer pays after N , then supplier charges the retailer an interest rate of I_{c2} /unit/year ($I_{c1} > I_{c2}$) on unpaid balance.

2. NOTATIONS:

- 2.1 P = the selling price / unit.
- 2.2 KD = the production rate per year, where $K > 1$
- 2.3 C = the unit purchase cost, with $C < P$.
- 2.4 M = the first offered credit period in settling the account without any charges.
- 2.5 N = the second permissible credit period in settling the account with interest charge I_{c2} on unpaid balance and $N > M$.
- 2.6 I_{c1} = the interest charged per \$ in stock per year by the supplier when retailer pays during $[M, N]$.
- 2.7 I_{c2} = the interest charged per \$ in stock per year by the supplier when retailer pays during $[N, T]$. ($I_{c1} > I_{c2}$)
- 2.8 I_e = the interest earned / \$ / year.
- 2.9 T = the replenishment cycle.
- 2.10 r = the discount rate ($r > \alpha$)
- 2.11 IE = the interest earned / time unit.
- 2.12 IC = the interest charged / time unit.

2.13 $(C_{vs} + \beta_2 t)$ = the setup cost for each production cycle for vendor.

2.14 $(C_{bs} + \beta_1 t)$ = the setup cost per order for buyer.

2.15 $(C_{hv} + \alpha_2 t)$ = holding cost per unit time for vendor.

2.16 $(C_{bh} + \alpha_1 t)$ = holding cost per unit time for buyer.

2.17 C_v = the unit cost for vendor.

2.18 C_b = the unit purchase cost for buyer.

2.19 S_b = shortage cost per unit time for buyer.

2.20 L_b = lost sale cost per unit time for buyer.

2.21 VC = the cost of vendor per unit time.

2.22 BC = the cost of buyer per unit time.

2.23 $TC(T)$ = total cost of an inventory system / time unit.

2.24 B = Backlogging rate.

2.25 The deterioration function

$$\theta(t, \delta) = \theta_0(\delta)t, \quad 0 < \theta_0(\delta) \ll 1, \quad t > 0$$

This is a special form of the two parameter weibull function considered by Covert and Philip. The function is some functions of the random variable α which range over a space and in which a p.d.f. $p(\delta)$ is defined such that

$$\int_{\Gamma} p(\delta) d\delta = 1$$

III. INDENTATIONS AND EQUATIONS MATHEMATICAL FORMULATION

The actual vendor's average inventory level in the integrated two-echelon inventory model is difference between the vendor's total average inventory level and the buyer's average inventory level. Since the inventory level is depleted due to a constant deterioration rate of the on-hand stock, the buyer's inventory level is represented by the following differential equation:

$$I'_b(t) + \theta_0(\delta)tI_b(t) = -\lambda_0 e^{\alpha t}, \quad 0 \leq t \leq t_1 \quad (1)$$

$$I'_b(t) = -B\lambda_0 e^{\alpha t}, \quad t_1 \leq t \leq T \quad (2)$$

The vendor's total inventory system consisting of production period and non-production period can be described as follows:

$$I'_{v1}(t) + \theta_0(\delta)tI_{v1}(t) = (K-1)\lambda_0 e^{\alpha t}, \quad 0 \leq t \leq T_1 \quad (3)$$

$$I'_{v2}(t) + \theta_0(\delta)tI_{v2}(t) = -\lambda_0 e^{\alpha t}, \quad 0 \leq t \leq T_2 \quad (4)$$

The boundary conditions are

$$I_{v1}(t) = 0, \quad t = 0 \quad (5)$$

$$I_{v2}(t) = 0, \quad t = T_2 \quad (6)$$

$$I_b(t) = I_0, \quad t = 0 \quad (7)$$

$$I_b(t) = 0, \quad t = t_1 \quad (8)$$

$$I_{v1}(T_1) = I_{v2}(0) \quad (9)$$

And,

$$T = \frac{T_2}{n} \quad (10)$$

The solutions of the above differential equations obtained are

$$I_b(t) = I_0 e^{-\frac{\theta_0(\delta)r^2}{2}} - \lambda_0 \left[t + \frac{\alpha t^2}{2} + \frac{\theta_0(\delta)t^3}{6} \right] e^{-\frac{\theta_0(\delta)r^2}{2}}, \quad 0 \leq t \leq t_1 \delta \quad (11)$$

$$I_b(t) = -\frac{B\lambda_0}{\alpha} \left[e^{\alpha t_1} - e^{\alpha t} \right], \quad t_1 \leq t \leq T \quad (12)$$

$$I_{v1}(t) = (K-1)\lambda_0 \left[t + \frac{\alpha t^2}{2} + \frac{\theta_0(\delta)t^3}{6} \right] e^{-\frac{\theta_0(\delta)r^2}{2}}, \quad 0 \leq t \leq T_1 \quad (13)$$

$$I_{v2}(t) = \lambda_0 \left[(T_2 - t) + \frac{\alpha}{2}(T_2^2 - t^2) + \frac{\theta_0(\delta)}{6}(T_2^3 - t^3) \right], \quad 0 \leq t \leq T_2 \quad (14)$$

Using the condition that one can get,

$$I_0 = -\lambda_0 \left[t_1 + \frac{\alpha t_1^2}{2} + \frac{\theta_0(\delta)t_1^3}{6} \right] \quad (15)$$

If the product of the deterioration rate and the replenishment interval is much smaller than one, the buyer's and the vendor's actual average inventory level, \bar{I}_b and \bar{I}_v , are

$$\begin{aligned} \bar{I}_b &= \frac{1}{T} \int_0^{t_1} e^{-rt} I_b(t) dt \\ &= \frac{I_0}{T} \left[t_1 - \frac{rt_1^2}{2} + (r^2 - \theta_0(\delta)) \frac{t_1^3}{6} \right] - \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} \right. \\ &\quad \left. + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{\theta_0(\delta)t_1^4}{12} - (3\alpha + 2r) \frac{\theta_0(\delta)t_1^5}{30} \right] \end{aligned} \quad (16)$$

and

$$\begin{aligned} \bar{I}_v &= \frac{1}{T_2} \left[\int_0^{T_1} e^{-rt} I_{v1}(t) dt + e^{-rT_1} \int_0^{T_2} e^{-rt} I_{v2}(t) dt \right] - \bar{I}_b \\ &= \frac{1}{T_2} \left[(K-1)\lambda_0 \left(\frac{T_1^2}{2} + \frac{T_1^3}{6}(\alpha - 2r) + \frac{T_1^4}{8}r(r - \alpha) + \frac{\alpha T_1^5 r^2}{20} - \frac{\theta_0(\delta)T_1^4}{12} - \frac{\theta_0(\delta)T_1^5(3\alpha + 2r)}{60} \right) \right. \\ &\quad \left. + e^{-rT_1} \lambda_0 \left(\frac{T_2^2}{2} - \frac{T_2^3}{6}(\alpha - 2r) + \frac{r^2 T_2^4}{24} - \frac{\alpha r T_2^4}{8} + \frac{\theta_0(\delta)T_2^4}{24} + \frac{\theta_0(\delta)T_2^5(4\alpha + 3r)}{60} \right) \right] - \bar{I}_b \end{aligned}$$

$$\begin{aligned} & - \frac{rt_1^2}{2} + (r^2 - \theta_0(\delta)) \frac{t_1^3}{6} + \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{\theta_0(\delta)t_1^4}{12} \right. \\ & \left. - (3\alpha + 2r) \frac{\theta_0(\delta)t_1^5}{30} \right] \end{aligned} \quad (17)$$

Respectively.

The annual total holding cost for the buyer and the vendor are

$$\begin{aligned} HC_b &= \frac{1}{T} \left[\int_0^{t_1} (C_{bh} + \alpha_1 t) e^{-rt} I_{b1}(t) dt \right] \\ &= -\frac{\lambda_0}{T} C_{bh} \left[\frac{t_1^2}{2} + \frac{t_1^3}{6}(\alpha - 2r) - \frac{t_1^4}{8}r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{\theta_0(\delta)t_1^4}{12} - \frac{\theta_0(\delta)t_1^5(3\alpha + 2r)}{60} \right] \\ &\quad - \frac{\alpha_1 \lambda_0}{T} \left[\frac{t_1^3}{3} + \frac{t_1^4}{8}(\alpha - 2r) - \frac{t_1^5}{10}r(r - \alpha) - \frac{\theta_0(\delta)t_1^5}{15} \right] + \frac{C_{bh} I_0}{T} \left[t_1 - \frac{rt_1^2}{2} \right. \\ &\quad \left. + \frac{(r^2 - \theta_0(\delta))t_1^3}{6} \right] + \frac{\alpha_1 I_0}{T} \left[\frac{t_1^2}{2} - \frac{rt_1^3}{3} + \frac{(r^2 - \theta_0(\delta))t_1^4}{8} \right] \end{aligned} \quad (18)$$

And

$$\begin{aligned} HC_v &= \frac{1}{T_2} \left[\int_0^{T_1} (C_{hv} + \alpha_2 t) e^{-rt} I_{v1}(t) dt + e^{-rT_1} \int_0^{T_2} (C_{hv} + \alpha_2 t) e^{-rt} I_{v2}(t) dt \right] - \bar{I}_b \\ &= \frac{1}{T_2} \left[(K-1)\lambda_0 C_{hv} \left(\frac{T_1^2}{2} + \frac{T_1^3}{6}(\alpha - 2r) - \frac{T_1^4}{8}r(r - \alpha) + \frac{\alpha T_1^5 r^2}{20} - \frac{\theta_0(\delta)T_1^4}{12} - \frac{\theta_0(\delta)T_1^5(3\alpha + 2r)}{60} \right) \right. \\ &\quad \left. + \frac{\alpha_2 (K-1)\lambda_0}{T_2} \left[\frac{T_1^3}{3} + \frac{T_1^4}{8}(\alpha - 2r) + \frac{T_1^5}{10}r(r - \alpha) - \frac{\theta_0(\delta)T_1^5}{15} \right] + \frac{C_{bh} \lambda_0 e^{-rT_1}}{T} \left[\frac{T_2^2}{2} - \frac{T_2^3(\alpha - 2r)}{6} \right. \right. \\ &\quad \left. \left. + \frac{r^2 T_2^4}{24} - \frac{\theta_0(\delta)T_2^4}{24} + \frac{\theta_0(\delta)T_2^5(4\alpha + 3r)}{60} \right] + \frac{\lambda_0 e^{-rT_1} \alpha_2}{T_2} \left[\frac{T_2^3}{6} + \frac{T_2^4(3\alpha - 2r)}{24} \right] - \frac{I_0}{T} \left[t_1 \right. \\ &\quad \left. - \frac{rt_1^2}{2} + (r^2 - \theta_0(\delta)) \frac{t_1^3}{6} + \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{\theta_0(\delta)t_1^4}{12} \right. \right. \\ &\quad \left. \left. - (3\alpha + 2r) \frac{\theta_0(\delta)t_1^5}{30} \right] \right] \end{aligned} \quad (19)$$

respectively.

The annual deterioration cost for the buyer and the vendor are

$$\begin{aligned} DC_b &= \frac{C_b}{T} \left[\int_0^{t_1} \theta_0(\delta) t e^{-rt} I_{b1}(t) dt \right] = \\ &= \frac{I_0 \theta_0(\delta)}{T} \left\{ \frac{t_1^2}{2} - \frac{rt_1^3}{6} + \frac{r^4 t_1^4}{8} \right\} - \frac{\lambda_0 \theta_0(\delta)}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4(\alpha - 2r)}{8} - \frac{t_1^5}{10}r(r - \alpha) \right\} \end{aligned} \quad (20)$$

and

$$DC_v = \frac{C_v}{T_2} \left[\int_0^{T_1} \theta_0(\delta) t e^{-rt} I_{v1}(t) dt + e^{-rT_1} \int_0^{T_2} \theta_0(\delta) t e^{-rt} I_{v2}(t) dt \right]$$

$$= C_v \left[\frac{(K-1)\lambda_0\theta_0(\delta)}{T_2} \left\{ \frac{T_1^3}{3} + \frac{T_1^4(\alpha-2r)}{8} + \frac{T_1^5 r(r-\alpha)}{10} \right\} + \frac{e^{-rT_1}\lambda_0\theta_0(\delta)}{T_2} \left\{ \frac{T_2^3}{6} + \frac{T_2^4(3\alpha-2r)}{24} \right\} \right] \quad (21)$$

respectively.

The annual set-up cost for the buyer and the vendor are

$$OC_b = \frac{1}{T} \left[\int_0^{t_1} (C_{bs} + \beta_1 t) dt + \int_{t_1}^T (C_{bs} + \beta_1 t) dt \right]$$

$$= C_{bs} + \frac{\beta_1 T}{2} \quad (22)$$

and

$$OC_v = \frac{1}{T_2} \left[\int_0^{T_1} (C_{vs} + \beta_2 t) dt + \int_0^{T_2} (C_{vs} + \beta_2 t) dt \right]$$

$$= [C_{vs}(T_1 + T_2) + \frac{\beta_2(T_1^2 + T_2^2)}{2}] \quad (23)$$

respectively.

The annual shortage cost for the buyer is

$$SC_b = \frac{S_b e^{-rt_1}}{T} \int_{t_1}^T \frac{\lambda_0}{\alpha} [e^{\alpha t_1} - e^{\alpha t}] e^{-rt} dt$$

$$= \frac{S_b \lambda_0 e^{-rt_1}}{T \alpha} \left[\frac{\alpha e^{(\alpha-r)t_1}}{r(\alpha-r)} - \frac{e^{(\alpha_1-rT)}}{r} - \frac{e^{(\alpha-r)T}}{(\alpha-r)} \right] \quad (24)$$

The annual lost sale cost for the buyer is

$$LC_b = \frac{L_b e^{-rt_1}}{T} \int_{t_1}^T (1-B) \lambda_0 e^{\alpha t} e^{-rt} dt$$

$$= \frac{L_b e^{-rt_1} \lambda_0 (1-B)}{T(\alpha-r)} [e^{(\alpha-r)T} - e^{(\alpha-r)t_1}] \quad (25)$$

The different costs associated with the system are set-up costs, holding costs, deterioration cost and shortage cost. Our aim is to minimize the total cost.

From (9), one can derive the following condition:

$$(K-1)\lambda_0 \left[T_1 + \frac{\alpha T_1^2}{2} + \frac{\theta_0(\delta) T_1^3}{6} \right] e^{-\frac{\theta_0(\delta) T_1^2}{2}} = \lambda_0 \left[T_2 + \frac{\alpha T_2^2}{2} + \frac{\theta_0(\delta) T_2^3}{6} \right] e^{-\frac{\theta_0(\delta) T_2^2}{2}} \quad (26)$$

By Taylor's series expansion, (4.26) is derived as

$$T_1 = \frac{1}{K-1} T_2 \left[1 + \frac{\alpha}{2} T_2 \right] \quad (27)$$

Regarding interest charged and interest earned based on the length of the cycle time t_1 , three cases arise:

IV. FIGURES AND TABLES

Regarding interest charged and interest earned based on the length of the cycle time t_1 , three cases arise:

Case I: $M \geq t_1$

Inventory level

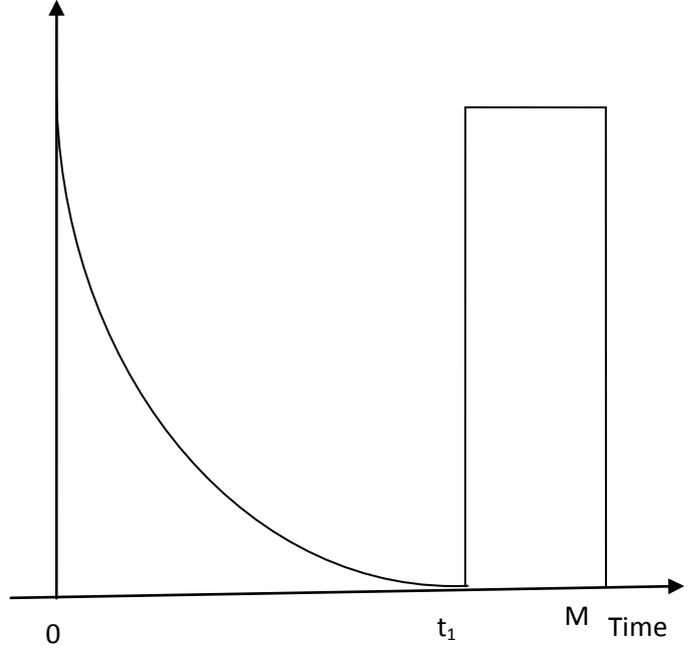


Fig 1: $t_1 \leq M$

In the first case, retailer does not pay any interest to the supplier. Here, retailer sells I_0 units during $(0, t_1)$ time interval and paying for CI_0 units in full to the supplier at time $M \geq t_1$, so interest charges are zero, i.e.

$$IC_1 = 0 \quad (28)$$

Retailers deposits the revenue in an interest bearing account at the rate of $Ie / \$ / \text{year}$. Therefore, interest earned IE_1 , per year is

$$IE_1 = \frac{PI_e}{T_2} \left[\int_0^{t_1} e^{-rt} D(t) t dt + (M - t_1) \int_0^{t_1} e^{-rt} D(t) dt \right]$$

$$= \frac{PI_e \lambda_0}{T_2} \left[(M - t_1) t_1 + \{1 + (\alpha - r)(M - t_1)\} \frac{t_1^2}{2} + \frac{t_1^3(\alpha - r)}{3} \right] \quad (29)$$

Total cost per unit time of an inventory system is

$$TC_b(t_1, \delta) = OC_b + HC_b + DC_b + SC_b + IC_1 - IE_1$$

$$= [C_{bs} + \frac{\beta_1 T}{2} - \frac{\lambda_0}{T} C_{bh} \left[\frac{t_1^2}{2} + \frac{t_1^3}{6} (\alpha - 2r) - \frac{t_1^4}{8} r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{\theta_0(\delta) t_1^4}{12} \right. \\ \left. - \frac{\theta_0(\delta) t_1^5 (3\alpha + 2r)}{60} \right] - \frac{\alpha_1 \lambda_0}{T} \left[\frac{t_1^3}{3} + \frac{t_1^4}{8} (\alpha - 2r) - \frac{t_1^5}{10} r(r - \alpha) - \frac{\theta_0(\delta) t_1^5}{15} \right]$$

$$\begin{aligned}
 & + \frac{C_{bh}I_0}{T} \left[t_1 - \frac{rt_1^2}{2} + \frac{(r^2 - \theta_0(\delta))t_1^3}{6} \right] + \frac{\alpha_1 I_0}{T} \left[\frac{t_1^2}{2} - \frac{rt_1^3}{3} + \frac{(r^2 - \theta_0(\delta))t_1^4}{8} \right] \\
 & + \left[\frac{I_0 \theta_0(\delta)}{T} \left\{ \frac{t_1^2}{2} - \frac{rt_1^3}{6} + \frac{r^4 t_1^4}{8} \right\} - \frac{\lambda_0 \theta_0(\delta)}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4(\alpha - 2r)}{8} - \frac{t_1^5}{10} r(r - \alpha) \right\} \right] \\
 & + \frac{S_b \lambda_0 e^{-rt_1}}{T\alpha} \left[\frac{\alpha e^{(\alpha-r)t_1}}{r(\alpha-r)} - \frac{e^{(\alpha_1-r)T}}{r} - \frac{e^{(\alpha-r)T}}{(\alpha-r)} \right] + \frac{L_b e^{-rt_1} \lambda_0 (1-B)}{T(\alpha-r)} [e^{(\alpha-r)T} - e^{(\alpha-r)t_1}] \\
 & - \frac{PI_e \lambda_0}{T_2} [(M - t_1)t_1 + \{1 + (\alpha - r)(M - t_1)\} \frac{t_1^2}{2} + \frac{t_1^3(\alpha - r)}{3}]
 \end{aligned} \tag{30}$$

Hence the mean cost

$$\langle TC_b \rangle = \int TC_b(t_1, \delta) p(\delta) d\delta \tag{31}$$

$$\begin{aligned}
 < TC_b > \\
 >= [C_{bs} + \frac{\beta_1 T}{2} - \frac{\lambda_0}{T} C_{bh} \left[\frac{t_1^2}{2} + \frac{t_1^3}{6} (\alpha - 2r) - \frac{t_1^4}{8} r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{A t_1^4}{12} \right. \\
 & \left. - \frac{A t_1^5 (3\alpha + 2r)}{60} \right] - \frac{\alpha_1 \lambda_0}{T} \left[\frac{t_1^3}{3} + \frac{t_1^4}{8} (\alpha - 2r) - \frac{t_1^5}{10} r(r - \alpha) - \frac{A t_1^5}{15} \right] \\
 & + \frac{C_{bh} I_0}{T} \left[t_1 - \frac{rt_1^2}{2} + \frac{(r^2 - A)t_1^3}{6} \right] + \frac{\alpha_1 I_0}{T} \left[\frac{t_1^2}{2} - \frac{rt_1^3}{3} + \frac{(r^2 - A)t_1^4}{8} \right] \\
 & + \left[\frac{I_0 A}{T} \left\{ \frac{t_1^2}{2} - \frac{rt_1^3}{6} + \frac{r^4 t_1^4}{8} \right\} - \frac{\lambda_0 A}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4(\alpha - 2r)}{8} - \frac{t_1^5}{10} r(r - \alpha) \right\} \right] \\
 & + \frac{S_b \lambda_0 e^{-rt_1}}{T\alpha} \left[\frac{\alpha e^{(\alpha-r)t_1}}{r(\alpha-r)} - \frac{e^{(\alpha_1-r)T}}{r} - \frac{e^{(\alpha-r)T}}{(\alpha-r)} \right] + \frac{L_b e^{-rt_1} \lambda_0 (1-B)}{T(\alpha-r)} [e^{(\alpha-r)T} - e^{(\alpha-r)t_1}] \\
 & - \frac{PI_e \lambda_0}{T_2} [(M - t_1)t_1 + \{1 + (\alpha - r)(M - t_1)\} \frac{t_1^2}{2} + \frac{t_1^3(\alpha - r)}{3}]
 \end{aligned} \tag{32}$$

Where $A = \int \theta_0(\delta) p(\delta) d\delta$ (33)

$$\langle TC_v \rangle = OC_v + HC_v + DC_v - IC_1$$

$$\begin{aligned}
 & = [C_{vs}(T_1 + T_2) + \frac{\beta_2(T_1^2 + T_2^2)}{2}] + \frac{1}{T_2} [(K-1)\lambda_0 C_m \left\{ \frac{T_1^2}{2} + \frac{T_1^3}{6} (\alpha - 2r) - \frac{T_1^4}{8} r(r - \alpha) \right\} \\
 & + \frac{\alpha T_1^5 r^2}{20} - \frac{A T_1^4}{12} - \frac{A T_1^5 (3\alpha + 2r)}{60}] + \frac{\alpha_2 (K-1)\lambda_0}{T_2} \left[\frac{T_1^3}{3} + \frac{T_1^4}{8} (\alpha - 2r) \right. \\
 & \left. + \frac{T_1^5}{10} r(r - \alpha) - \frac{A T_1^5}{15} \right] + \frac{C_{bh} \lambda_0 e^{-rt_1}}{T} \left[\frac{T_2^2}{2} - \frac{T_2^3 (\alpha - 2r)}{6} + \frac{r^2 T_2^4}{24} - \frac{A T_2^4}{24} \right. \\
 & \left. + \frac{A T_2^5 (4\alpha + 3r)}{60} \right] + \frac{\lambda_0 e^{-rt_1} \alpha_2}{T_2} \left[\frac{T_2^3}{6} + \frac{T_2^4 (3\alpha - 2r)}{24} \right] - \frac{I_0}{T} \left[t_1 - \frac{rt_1^2}{2} \right. \\
 & \left. + (r^2 - A) \frac{t_1^3}{6} \right] + \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{A t_1^4}{12} \right]
 \end{aligned}$$

$$\begin{aligned}
 & - (3\alpha + 2r) \frac{A t_1^5}{30} + C_v \left[\frac{(K-1)\lambda_0 A}{T_2} \left\{ \frac{T_1^3}{3} + \frac{T_1^4 (\alpha - 2r)}{8} + \frac{T_1^5 r(r - \alpha)}{10} \right\} \right. \\
 & \left. + \frac{e^{-rt_1} \lambda_0 A}{T_2} \left\{ \frac{T_2^3}{6} + \frac{T_2^4 (3\alpha - 2r)}{24} \right\} \right]
 \end{aligned} \tag{34}$$

To minimize the total cost per unit time, the optimum value of t_1, T_2 is the solution of following equation.

Case II: $M < t_1 < N$

Inventory level

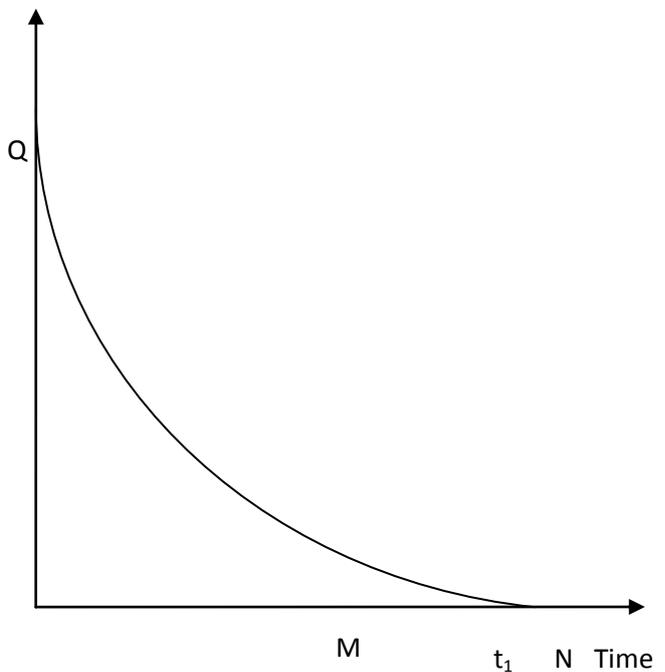


Fig: 2 $M < t_1 < N$

In the second case, supplier charges interest at the rate IC_1 on unpaid balance.

Interest earned, IE_2 during $[0, M]$ is

$$\begin{aligned}
 IE_2 & = PI_e \int_0^M e^{-rt} D(t) dt \\
 & = PI_e \lambda_0 \left[\frac{M^2}{2} + \frac{M^3}{6} (\alpha - r) + \frac{M^4}{8} (\alpha - r)^2 \right]
 \end{aligned} \tag{35}$$

Retailer pay for I_0 units purchased at time $t = 0$ at the rate of $C / \$ /$ unit to the supplier during $[0, M]$. The retailer sells $D(M).M$ units at selling price $P /$ unit. So, he has generated revenue of $P D(M).M + IE_2$. Then two sub cases may arise:

Sub Case: 2.1

Let $P D(M).M + IE_2 \geq CI_0$, i.e. retailer has enough money to settle his account for all I_0 units procured at time $t = 0$. Then interest charge will be

$$IC_{2,1} = 0 \tag{36}$$

and interest earned

$$IE_{2.1} = \frac{IE_2}{T_2}$$

$$= \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (37)$$

So, total cost $TC_{2.1}$ per unit time of inventory system is

$$\langle TC_b \rangle = OC_b + HC_b + DC_b + SC_b + LC_b + IC_{2.1} - IE_{2.1}$$

$$S = [C_{bs} + \frac{\beta_1 T}{2} - \frac{\lambda_0}{T} C_{bh} \left\{ \frac{t_1^2}{2} + \frac{t_1^3}{6}(\alpha - 2r) - \frac{t_1^4}{8}r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{At_1^4}{12} \right. \\ \left. - \frac{At_1^5(3\alpha + 2r)}{60} \right\} - \frac{\alpha_1 \lambda_0}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4}{8}(\alpha - 2r) - \frac{t_1^5}{10}r(r - \alpha) - \frac{At_1^5}{15} \right\} \\ + \frac{C_{bh} I_0}{T} \left[t_1 - \frac{rt_1^2}{2} + \frac{(r^2 - A)t_1^3}{6} \right] + \frac{\alpha_1 I_0}{T} \left[\frac{t_1^2}{2} - \frac{rt_1^3}{3} + \frac{(r^2 - A)t_1^4}{8} \right] \\ + \left[\frac{I_0 A}{T} \left\{ \frac{t_1^2}{2} - \frac{rt_1^3}{6} + \frac{r^4 t_1^4}{8} \right\} - \frac{\lambda_0 A}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4(\alpha - 2r)}{8} - \frac{t_1^5}{10}r(r - \alpha) \right\} \right] \\ + \frac{S_b \lambda_0 e^{-rt_1}}{T\alpha} \left[\frac{\alpha e^{(\alpha-r)t_1}}{r(\alpha-r)} - \frac{e^{(\alpha_1-r)T}}{r} - \frac{e^{(\alpha-r)T}}{(\alpha-r)} \right] + \frac{L_b e^{-rt_1} \lambda_0 (1-B)}{T(\alpha-r)} [e^{(\alpha-r)T} - e^{(\alpha-r)t_1}] \\ - \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (38)$$

$$\langle TC_v \rangle = OC_v + HC_v + DC_v - IC_{2.1}$$

$$= [C_{vs}(T_1 + T_2) + \frac{\beta_2(T_1^2 + T_2^2)}{2}] + \frac{1}{T_2} [(K-1)\lambda_0 C_{bv} \left\{ \frac{T_1^2}{2} + \frac{T_1^3}{6}(\alpha - 2r) - \frac{T_1^4}{8}r(r - \alpha) \right. \\ \left. + \frac{\alpha T_1^5 r^2}{20} - \frac{AT_1^4}{12} - \frac{AT_1^5(3\alpha + 2r)}{60} \right\}] + \frac{\alpha_2(K-1)\lambda_0}{T_2} \left[\frac{T_1^3}{3} + \frac{T_1^4}{8}(\alpha - 2r) \right] \\ + \frac{T_1^5}{10}r(r - \alpha) - \frac{AT_1^5}{15}] + \frac{C_{bv} \lambda_0 e^{-rt_1}}{T} \left[\frac{T_2^2}{2} - \frac{T_2^3(\alpha - 2r)}{6} + \frac{r^2 T_2^4}{24} - \frac{AT_2^4}{24} \right. \\ \left. + \frac{AT_2^5(4\alpha + 3r)}{60} \right] + \frac{\lambda_0 e^{-rt_1} \alpha_2}{T_2} \left[\frac{T_2^3}{6} + \frac{T_2^4(3\alpha - 2r)}{24} \right] - \frac{I_0}{T} \left[t_1 - \frac{rt_1^2}{2} \right. \\ \left. + (r^2 - A)\frac{t_1^3}{6} \right] + \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r)\frac{t_1^3}{6} + r(r - \alpha)\frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{At_1^4}{12} \right. \\ \left. - (3\alpha + 2r)\frac{At_1^5}{30} \right] + C_v \left[\frac{(K-1)\lambda_0 A}{T_2} \left\{ \frac{T_1^3}{3} + \frac{T_1^4(\alpha - 2r)}{8} + \frac{T_1^5 r(r - \alpha)}{10} \right\} \right. \\ \left. + \frac{e^{-rt_1} \lambda_0 A}{T_2} \left\{ \frac{T_2^3}{6} + \frac{T_2^4(3\alpha - 2r)}{24} \right\} \right] - \\ \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (39)$$

To minimize the total cost per unit time, the optimum value of t_1, T_2 is the solution of following equation.

Sub Case: 2.2

Let $P D(M).M + IE_2 < CI_0$. Here, retailer will have to pay interest on unpaid balance $U_1 = CI_0 - (P D(M).M + IE_2)$ at the rate of Ic_1 at time M to the supplier. Then interest paid per unit time is given by

$$IC_{2.2} = \frac{U_1^2 Ic_1}{PI_0} \int_M^{t_1} e^{-rt} I(t) dt$$

$$= \frac{U_1^2 Ic_1}{PI_0} \left[\frac{(t_1^2 - M^2)}{2} + \frac{(\alpha - r)(t_1^3 - M^3)}{2} + \frac{(\alpha - r)^2(t_1^4 - M^4)}{8} \right] \quad (40)$$

Where,

$$U_1 = CI_0 - (P D(M).M + IE_2)$$

$$= CI_0 - P\lambda_0 \left[M + \left(\alpha + \frac{I_e}{2} \right) M^2 + \left(\frac{\alpha^2}{2} + \frac{I_e(\alpha - r)}{2} \right) M^3 + \left(\frac{\alpha^3}{6} + \frac{I_e(\alpha - r)^2}{8} \right) M^4 \right] \quad (41)$$

And interest earned

$$IE_{2.2} = \frac{IE_2}{T_2}$$

$$= \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (42)$$

So, total cost $TC_{2.2}$ per unit time of inventory system is

$$\langle TC_b \rangle = OC_b + HC_b + DC_b + SC_b + LC_b + IC_{2.2} - IE_{2.2}$$

$$= [C_{bs} + \frac{\beta_1 T}{2} - \frac{\lambda_0}{T} C_{bh} \left\{ \frac{t_1^2}{2} + \frac{t_1^3}{6}(\alpha - 2r) - \frac{t_1^4}{8}r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{At_1^4}{12} \right. \\ \left. - \frac{At_1^5(3\alpha + 2r)}{60} \right\} - \frac{\alpha_1 \lambda_0}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4}{8}(\alpha - 2r) - \frac{t_1^5}{10}r(r - \alpha) - \frac{At_1^5}{15} \right\} \\ + \frac{C_{bh} I_0}{T} \left[t_1 - \frac{rt_1^2}{2} + \frac{(r^2 - A)t_1^3}{6} \right] + \frac{\alpha_1 I_0}{T} \left[\frac{t_1^2}{2} - \frac{rt_1^3}{3} + \frac{(r^2 - A)t_1^4}{8} \right] \\ + \left[\frac{I_0 A}{T} \left\{ \frac{t_1^2}{2} - \frac{rt_1^3}{6} + \frac{r^4 t_1^4}{8} \right\} - \frac{\lambda_0 A}{T} \left\{ \frac{t_1^3}{3} + \frac{t_1^4(\alpha - 2r)}{8} - \frac{t_1^5}{10}r(r - \alpha) \right\} \right] \\ + \frac{S_b \lambda_0 e^{-rt_1}}{T\alpha} \left[\frac{\alpha e^{(\alpha-r)t_1}}{r(\alpha-r)} - \frac{e^{(\alpha_1-r)T}}{r} - \frac{e^{(\alpha-r)T}}{(\alpha-r)} \right] + \frac{L_b e^{-rt_1} \lambda_0 (1-B)}{T(\alpha-r)} [e^{(\alpha-r)T} - e^{(\alpha-r)t_1}] \\ - \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (43)$$

$$\langle TC_v \rangle = OC_v + HC_v + DC_v - IC_{2.2}$$

$$= S [C_{vs}(T_1 + T_2) + \frac{\beta_2(T_1^2 + T_2^2)}{2}] + \frac{1}{T_2} [(K-1)\lambda_0 C_{bv} \left\{ \frac{T_1^2}{2} + \frac{T_1^3}{6}(\alpha - 2r) - \frac{T_1^4}{8}r(r - \alpha) \right.$$

$$\begin{aligned}
 & + \frac{\alpha T_1^5 r^2}{20} - \frac{AT_1^4}{12} - \frac{AT_1^5(3\alpha + 2r)}{60} \Big] + \frac{\alpha_2(K-1)\lambda_0}{T_2} \left[\frac{T_1^3}{3} + \frac{T_1^4}{8}(\alpha - 2r) \right. \\
 & + \frac{T_1^5}{10}r(r - \alpha) - \frac{AT_1^5}{15} \Big] + \frac{C_{bh}\lambda_0 e^{-rT_1}}{T} \left[\frac{T_2^2}{2} - \frac{T_2^3(\alpha - 2r)}{6} + \frac{r^2 T_2^4}{24} - \frac{AT_2^4}{24} \right. \\
 & + \frac{AT_2^5(4\alpha + 3r)}{60} \Big] + \frac{\lambda_0 e^{-rT_1} \alpha_2}{T_2} \left[\frac{T_2^3}{6} + \frac{T_2^4(3\alpha - 2r)}{24} \right] - \frac{I_0}{T} \left[t_1 - \frac{rt_1^2}{2} \right. \\
 & + (r^2 - A) \frac{t_1^3}{6} \Big] + \frac{\lambda_0}{T} \left[\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{At_1^4}{12} \right. \\
 & \left. - (3\alpha + 2r) \frac{At_1^5}{30} \right] + C_v \left[\frac{(K-1)\lambda_0 A}{T_2} \left\{ \frac{T_1^3}{3} + \frac{T_1^4(\alpha - 2r)}{8} + \frac{T_1^5 r(r - \alpha)}{10} \right\} \right. \\
 & + \frac{e^{-rT_1} \lambda_0 A}{T_2} \left\{ \frac{T_2^3}{6} + \frac{T_2^4(3\alpha - 2r)}{24} \right\} \Big] - \frac{U_1^2 I_{c1}}{PI_0} \left[\frac{(t_1^2 - M^2)}{2} \right. \\
 & \left. + \frac{(\alpha - r)(t_1^3 - M^3)}{2} + \frac{(\alpha - r)^2(t_1^4 - M^4)}{8} \right] \quad (44)
 \end{aligned}$$

To minimize the total cost per unit time, the optimum value of t_1, T_2 is the solution of following equation.

Case III: $t_1 \square N$

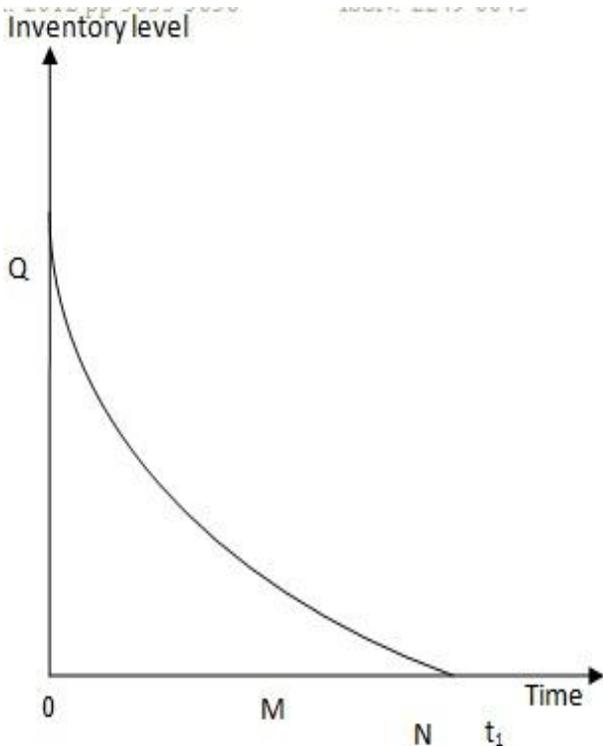


Fig 3: $t_1 \square N$

In the final case, retailer pays interest at the rate of I_{c2} to the supplier. Based on the total purchased cost, CI_0 , total money $P D(M).M + IE_2$

$$= \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right]$$

in account at M and total money

$PD(N).N = P \lambda_0 e^{-\lambda N} N$ at N, there are three sub cases may arise:

Sub Case 3.1 Let $P D(M).M + IE_2 \geq CI_0$

This case is same as sub case 2.1, here 3.1 designate decision variables and objective function.

Sub Case 3.2 Let $P D(M).M + IE_2 < CI_0$ and

$$PD(N - M).(N - M) + PI_e \int_M^N D(t) dt \geq CI_0 - (PD(M).M + IE_2)$$

$$P \lambda_0 e^{\alpha(N-M)} (N - M) + PI_e \lambda_0 [(N - M) + \frac{\alpha}{2}(N^2 - M^2)] \geq CI_0 - (P \lambda_0 e^{\alpha M} M + IE_2)$$

This case similar to sub case 2.2.

Sub Case 3.3 Let $P D(M).M + IE_2 < CI_0$ and

$$P \lambda_0 e^{\alpha(N-M)} (N - M) + PI_e \lambda_0 [(N - M) + \frac{\alpha}{2}(N^2 - M^2)] < CI_0 - (P \lambda_0 e^{\alpha M} M + IE_2)$$

Here, retailer does not have enough money to pay off total purchase cost at N. He will not pay money of $P D(M).M + IE_2$ at M and $PD(N - M).(N - M) + PI_e [(N - M) + \frac{\alpha}{2}(N^2 - M^2)]$ at

N. That's why he has to pay interest on unpaid balance $U_1 = CI_0 - (P D(M).M + IE_2)$ with I_{c1} interest rate during (M, N)

and $U_2 = U_1 - PD(N - M).(N - M) + PI_e \int_M^N D(t) dt$

with interest rate I_{c2} during (N, t_1).

Therefore, total interest charged on retailer, $IC_{3.3}$ per unit time is

$$\begin{aligned}
 IC_{3.3} &= \frac{U_1 I_{c1} (N - M)}{T_2} + \frac{U_2^2 I_{c1}}{PI_0} \int_N^{t_1} e^{-rt} I_b(t) dt \\
 &= \frac{U_1 I_{c1} (N - M)}{T_2} + \frac{U_2^2 I_{c1}}{PI_0} \left[\frac{t_1^2}{2} - N^2 + \frac{(t_1^3 - N^3)(\alpha - 2r)}{6} + \frac{(t_1^4 - N^4)}{8} r(r - \alpha) \right. \\
 &+ \frac{(t_1^5 - N^5)}{20} r^2 \alpha - \frac{\theta_0(\delta)(t_1^4 - N^4)}{12} - \frac{\theta_0(\delta)(t_1^5 - N^5)(3\alpha + 2r)}{12} \Big] - \frac{U_2^2 I_{c1}}{P} [(t_1 - N) \\
 &\left. - \frac{r(t_1^2 - N^2)}{2} + \frac{(r^2 - \theta_0(\delta))(t_1^3 - N^3)}{6} \right] \quad (45)
 \end{aligned}$$

Interest earned per unit time is

$$\begin{aligned}
 IE_{3.3} &= \frac{IE_2}{T_2} \\
 &= \frac{PI_e \lambda_0}{T_2} \left[\frac{M^2}{2} + \frac{M^3}{6}(\alpha - r) + \frac{M^4}{8}(\alpha - r)^2 \right] \quad (46)
 \end{aligned}$$

So, total cost $TC_{3.3}$ per unit time of inventory system is

$$<TC_b > = OC_b + HC_b + DC_b + SC_b + LC_b + IC_{3.3} - IE_{3.3}$$

$$\begin{aligned}
 &= [C_{bs} + \frac{\beta_1 T}{2} - \frac{\lambda_0}{T} C_{bh} [\frac{t_1^2}{2} + \frac{t_1^3}{6} (\alpha - 2r) - \frac{t_1^4}{8} r(r - \alpha) + \frac{\alpha t_1^5 r^2}{20} - \frac{A t_1^4}{12} \\
 &- \frac{A t_1^5 (3\alpha + 2r)}{60}] - \frac{\alpha_1 \lambda_0}{T} [\frac{t_1^3}{3} + \frac{t_1^4}{8} (\alpha - 2r) - \frac{t_1^5}{10} r(r - \alpha) - \frac{A t_1^5}{15}] \\
 &+ \frac{C_{bh} I_0}{T} [t_1 - \frac{r t_1^2}{2} + \frac{(r^2 - A) t_1^3}{6}] + \frac{\alpha_1 I_0}{T} [\frac{t_1^2}{2} - \frac{r t_1^3}{3} + \frac{(r^2 - A) t_1^4}{8} \\
 &+ [\frac{I_0 A}{T} \{\frac{t_1^2}{2} - \frac{r t_1^3}{6} + \frac{r^4 t_1^4}{8}\} - \frac{\lambda_0 A}{T} \{\frac{t_1^3}{3} + \frac{t_1^4 (\alpha - 2r)}{8} - \frac{t_1^5}{10} r(r - \alpha)\}] \\
 &+ \frac{S_b \lambda_0 e^{-r t_1}}{T \alpha} [\frac{\alpha e^{(\alpha - r) t_1}}{r(\alpha - r)} - \frac{e^{(\alpha_1 - r) T}}{r} - \frac{e^{(\alpha - r) T}}{(\alpha - r)}] + \frac{L_b e^{-r t_1} \lambda_0 (1 - B)}{T(\alpha - r)} [e^{(\alpha - r) T} - e^{(\alpha_1 - r) t_1}] \\
 &+ \frac{U_1 I_c (N - M)}{T_2} + \frac{U_2^2 I_{c1}}{P I_0} [\frac{t_1^2}{2} - N^2 + \frac{(t_1^3 - N^3)(\alpha - 2r)}{6} + \frac{(t_1^4 - N^4)}{8} r(r - \alpha) \\
 &+ \frac{(t_1^5 - N^5)}{20} r^2 \alpha - \frac{\theta_0 (\delta) (t_1^4 - N^4)}{12} - \frac{\theta_0 (\delta) (t_1^5 - N^5) (3\alpha + 2r)}{12}] - \frac{U_2^2 I_{c1}}{P} [(t_1 - N) \\
 &- \frac{r(t_1^2 - N^2)}{2} + \frac{(r^2 - \theta_0 (\delta))(t_1^3 - N^3)}{6}] \\
 &- \frac{P I_e \lambda_0}{T_2} [\frac{M^2}{2} + \frac{M^3}{6} (\alpha - r) + \frac{M^4}{8} (\alpha - r)^2] \quad (47)
 \end{aligned}$$

$$TC_v = OC_v + HC_v + DC_v - IC_{3.3}$$

$$\begin{aligned}
 &= [C_{vs} (T_1 + T_2) + \frac{\beta_2 (T_1^2 + T_2^2)}{2}] + \frac{1}{T_2} [(K - 1) \lambda_0 C_m \{\frac{T_1^2}{2} + \frac{T_1^3}{6} (\alpha - 2r) - \frac{T_1^4}{8} r(r - \alpha) \\
 &+ \frac{\alpha T_1^5 r^2}{20} - \frac{A T_1^4}{12} - \frac{A T_1^5 (3\alpha + 2r)}{60}\}] + \frac{\alpha_2 (K - 1) \lambda_0}{T_2} [\frac{T_1^3}{3} + \frac{T_1^4}{8} (\alpha - 2r) \\
 &+ \frac{T_1^5}{10} r(r - \alpha) - \frac{A T_1^5}{15}] + \frac{C_{bh} \lambda_0 e^{-r T_1}}{T} [\frac{T_2^2}{2} - \frac{T_2^3 (\alpha - 2r)}{6} + \frac{r^2 T_2^4}{24} - \frac{A T_2^4}{24} \\
 &+ \frac{A T_2^5 (4\alpha + 3r)}{60}] + \frac{\lambda_0 e^{-r T_1} \alpha_2}{T_2} [\frac{T_2^3}{6} + \frac{T_2^4 (3\alpha - 2r)}{24}] - \frac{I_0}{T} [t_1 - \frac{r t_1^2}{2} \\
 &+ (r^2 - A) \frac{t_1^3}{6}] + \frac{\lambda_0}{T} [\frac{t_1^2}{2} + (\alpha - 2r) \frac{t_1^3}{6} + r(r - \alpha) \frac{t_1^4}{8} + \frac{\alpha r^2 t_1^5}{20} - \frac{A t_1^4}{12} \\
 &- (3\alpha + 2r) \frac{A t_1^5}{30}] + C_v [\frac{(K - 1) \lambda_0 A}{T_2} \{\frac{T_1^3}{3} + \frac{T_1^4 (\alpha - 2r)}{8} + \frac{T_1^5 r(r - \alpha)}{10}\} \\
 &+ \frac{e^{-r T_1} \lambda_0 A}{T_2} \{\frac{T_2^3}{6} + \frac{T_2^4 (3\alpha - 2r)}{24}\}] - \\
 &\frac{P I_e \lambda_0}{T_2} [\frac{M^2}{2} + \frac{M^3}{6} (\alpha - r) + \frac{M^4}{8} (\alpha - r)^2] \quad (48)
 \end{aligned}$$

To minimize the total cost per unit time, the optimum value of t_1, T_2 is the solution of following equation.

NUMERICAL ILLUSTRATION: THE PRECEDING THEORY CAN BE ILLUSTRATED BY THE FOLLOWING NUMERICAL EXAMPLE WHERE THE PARAMETERS ARE GIVEN AS FOLLOWS:

Demand parameters, $a = 500, b = 5, c = 2$

Selling price, $P = 30$

Buyer's purchased cost, $C_b = 35$

Buyer's percentage holding cost per year per dollar,

$C_{bh} = 0.2$

Buyer's ordering cost per order, $C_{bs} = 500$

Buyer's shortage cost, $S_b = 50$

Vendor's unit cost, $C_v = 20$

Vendor's percentage holding cost per year per dollar,

$C_{vh} = 0.2$

Vendor's setup cost per order, $C_{vs} = 1000$

Vendor's production rate per year, $K = 5$

Deterioration rate, $\theta_0 (\delta) = 0.01$

First delay period, $M = 0.2$

Second delay period, $N = 0.4$

The interest earned, $I_e = 0.05$

The interest charged, $I_{c1} = 0.10$

The interest charged, $I_{c2} = 0.20$ ($I_{c1} > I_{c2}$)

Backlogging rate, $B = 0$

Table 1:

N	T ₂	t ₁	VC	BC	TC
1	0.827183	0.800625	1757.09	1405.95	3163.03
2	0.942755	0.456282	2086.02	1517.28	3603.30
3	1.02889	0.331991	2274.14	1790.02	4064.16
4	1.10067	0.266369	2425.17	2083.54	4508.71
5	1.16312	0.225188	2559.84	2374.64	4935.49

Table 2:

N	T ₂	t ₁	VC	BC	TC
1	0.792393	0.745431	1966.30	1774.98	3741.28
2	0.921355	0.433375	2435.64	1877.95	4313.59
3	1.01214	0.317385	2708.53	1969.85	4678.39
4	1.08612	0.255438	2927.92	2215.75	5143.68
5	1.14978	0.216326	3121.89	2474.02	5595.9

Table 3:

N	T ₂	t ₁	VC	BC	TC
1	0.792393	0.745431	1780.22	1823.14	3603.36
2	0.921355	0.433375	1934.79	1957.21	3892.00
3	1.01214	0.317385	2265.29	2049.26	4314.55
4	1.08612	0.255438	2315.26	2320.36	4635.62
5	1.14978	0.216326	2497.69	2546.68	5044.37

Table 4:

N	T ₂	t ₁	VC	BC	TC
1	1.43526	0.40970	1524.28	6918.29	8442.57
2	2.22410	0.41285	1328.68	4504.47	5833.15
3	2.69166	0.416398	1270.01	3076.3	4346.31
4	3.01918	0.420088	1172.2	2283.4	3455.6
5	3.27318	0.423883	1032.2	1856.66	2888.86

V. CONCLUSION

Here we have studied a two echelon supply chain with some very realistic assumptions. We studied our model in a progressive credit period. No doubt, this assumption imparts an economic viability to the whole study. In real world, it is noted that, as a result of progressive permissible delay in settling the replenishment account, the economic replenishment interval and order quantity generally increase marginally, although the annual cost decreases considerably. The saving in cost as a result of permissible delay in settling the replenishment account largely come the ability to delay payment without paying any interest. As a result of increasing order quantity under conditions or permissible delay in payments, we need to order less often. So this EOQ model is applicable when supplier gives the trade credit to the retailer.

AKNOWLEDGEMENT

We are thankful to Dr.A.P.Singh of SGRR (PG) College Dehradun for his constant support and review of our research paper. We also thank Prof. Yudhveer Singh Moudgil of Uttranchal Institute of Technology for reviewing our research paper.

OBSERVATION

The data obtained clearly shows that individual optimal solutions are very different from each other. However, there exists a solution which ultimately provides the minimum operating cost to the whole supply chain. All the observations can be summed up as follows:

1. An increase in the interest charged, increases the buyer cost BC and decrease the vendor cost VC of the commodity.
2. Optimal solution for the buyer is n=1 in table first while for the vendor, it is n=5 in table 4. The overall optimal solution which ultimately minimizes the cost across the whole supply chain is n=5 in table 4

REFERENCES

- 1) Hwang H. S. (1997). A Study on an inventory model for items with Weibull ameliorating. *Com. & Ind. Eng.* **33**, 701-704.
- 2) Hwang H. S. (1999). Inventory Model for Items for both Deteriorating and Ameliorating Items. *Com. & Ind. Eng.* **37**, 257-260.
- 3) Hwang, H. S. (2004): A Stochastic Set-covering Location Model for Both Ameliorating and Deteriorating Items. *Com. & Ind. Eng.* **46**, 313-319.
- 4) Law S. T. and Wee H. M. (2006). An integrated production-inventory model for ameliorating and deteriorating items taking account of time discounting. *Math. & Comp. Mod.*, **43**, 673-685.
- 5) Mondal B., Bhunia A. K. and Maiti M. (2003). An inventory system of ameliorating items for price dependent demand rate. *Com. & Ind. Eng.* **45**, 443-456.
- 6) Moon I., Giri B. C. and Ko B. (2005). Economic order quantity models for ameliorating/ deteriorating items under inflation and time discounting. *Euro. Jour. Oper. Res.* **162**, 773-785.

Inventory Model: Deteriorating Items with Price and Time Dependent Demand Rate

Jasvinder Kaur¹, Rajendra Sharma²

*(Department Of Mathematics, Graphic Era University Dehradun, India)

** (Department Of Mathematics Graphic Era University Dehradun, india)

ABSTRACT: This study presents a deterministic inventory model for deteriorating products under the condition of instantaneous replenishment. The rate of deterioration is assumed to be a constant fraction of on hand inventory and demand is a function of selling price and decreases exponentially with time. It is shown that the developed model can be related to Cohen Model and Standard model without deterioration. A numerical example demonstrates the effectiveness of the developed model.

Keywords: Demand; deterioration; inventory; optimal; shortage.

I. INTRODUCTION

It is of considerable interest and important to analysis the inventory models for deteriorating items. In many inventory systems, the effect of deterioration is an important factor and cannot be ignored. Deterioration may be defined as decay or damage or spoilage, so that the item can not be used as in its original state. Thus for example, blood, certain food items, photographic films, fruits, chemicals, radio active substances are some examples of items in which deterioration plays a major role.

Ghare and Schrader (Ghare and Schrader, 1963) developed a model for exponentially decaying inventory considering constant demand. Emmons (1968) also developed a model for exponential decaying products, where the product decayed at one rate into a new product, which decayed at a second rate.

Cohen (1977) developed a model for joint pricing and ordering policy for exponentially decaying inventory with known demand and constant decay rate. Mukherjee (1987) extended it by considering time varying decay rate. Kumar and Sharma (2009) has extended Mukherjee's (1999) model by considering shortages.

In this study, we consider demand as a function of time and price both and decreases exponentially. Deterioration is assumed to be constant fraction of on hand inventory. More ever the developed model reduces Cohen's model (1977) and standard model without deterioration.

II. PROPOSED ASSUMPTIONS & NOTATIONS

The model is developed under following assumptions and notations.

- 1 Deterioration rate λ is a constant fraction of on hand inventory.
- 2 Demand rate $D(t, p)$ is known and decreases exponentially, i.e. at time $t, t \geq 0$

$$D(t, p) = \frac{e^{-\theta t}}{d(p)}$$

θ is constant governing the decreasing rate of demand. p is the selling price per unit and $d(p)$ is the function of p .

- 3 Lead time is zero, shortages are not allowed.
- 4 The replenishment rate is infinite and T is the cycle time.
- 5 There is no replacement or repair of the decayed units during the period under consideration.
- 6 The unit purchase cost is C, h is the holding cost per unit per unit time, K is the ordering cost per order.
- 7 $I(t)$ is the inventory at any time t .

III. MATHEMATICAL MODELLING AND ANALYSIS

The differential equation governing the system is given by

$$\frac{dI(t)}{dt} = -\lambda I(t) - D(p, t) \quad (1)$$

Solution of (1) is given by after adjusting constant of integration.

$$I(t) = I(0)e^{-\lambda t} + \frac{e^{-\theta t} - e^{-\lambda t}}{d(p)(\theta - \lambda)} \quad (2)$$

Inventory without decay at time t is given by the differential equation.

$$\frac{d}{dt} I_W(t) = -\frac{e^{-\theta t}}{d(p)} \quad (3)$$

The solution of which gives

$$I_W(t) = I(0) + \frac{(e^{-\theta t} - 1)}{\theta.d(p)} \quad (4)$$

The stock loss $Z(t)$ due to decay in $[0, T]$ is given by

$$\begin{aligned} z(t) &= I_W(t) - I(t) \\ &= I(0) + \frac{e^{-\theta t} - 1}{\theta.d(p)} - I(t) \end{aligned} \quad (5)$$

Using (2), equation (5) reduced to

$$Z(t) = I(t)(e^{\lambda t} - 1) - \frac{e^{-(\theta-\lambda)t} - 1}{d(p)(\theta - \lambda)} + \frac{e^{-\theta t} - 1}{\theta.d(p)} \quad (6)$$

Total demand D during $(0, T)$ is given by

$$D = \int_0^T \frac{e^{-\theta t}}{d(p)} dt = \frac{-(e^{-\theta T} - 1)}{\theta.d(p)} \quad (7)$$

Also $Z(T) = \frac{e^{-\theta T} - 1}{\theta d(p)} - \frac{e^{-(\theta-\lambda)T} - 1}{d(p)(\theta-\lambda)}$ (8)

Lot Size $Q_T = \frac{1 - e^{-(\theta-\lambda)T}}{d(p)(\theta-\lambda)}$ (9)

Also $I(0) = Q_T$, then Using (9), (2) reduces to

$I(t) = \frac{e^{-\theta t} - e^{-\theta T} \cdot e^{-\lambda(t-T)}}{d(p)(\theta-\lambda)}$ (10)

Cost per cycle becomes

$C^*(T, p) = K + C \cdot Q_T + h \int_0^T I(t) dt$

For a fixed price level p cost per unit time $C(T, p)$ is

$C(T, p) = C^*(T, p)/T$

$$= \frac{K}{T} + \frac{C[1 - e^{-(\theta-\lambda)T}]}{Td(p)(\theta-\lambda)} + \frac{h}{Td(p)(\theta-\lambda)} \left[\frac{1 - e^{-\theta T}}{\theta} - \frac{e^{-(\theta-\lambda)T}(1 - e^{-\lambda T})}{\lambda} \right]$$
 (11)

By holding p fixed, the necessary conditions for minimizing $C(T, p)$ with respect to T is

$\frac{\partial C(T, p)}{\partial T} = 0$

implies

$$\frac{1}{T^2} \left[-K - \frac{C}{d(p)(\theta-\lambda)} + \frac{(C\lambda + h)e^{-(\theta-\lambda)T} [T(\theta-\lambda) + 1]}{\lambda d(p)(\theta-\lambda)} - \frac{h\theta T^2}{2d(p)(\theta-\lambda)} - \frac{h}{\lambda d(p)(\theta-\lambda)} \right] = 0$$

$$\Rightarrow \frac{(C\lambda + h)e^{-(\theta-\lambda)T} [T(\theta-\lambda) + 1]}{\lambda d(p)(\theta-\lambda)} + \frac{h\theta T^2}{2\lambda d(p)} = K + \frac{C\lambda + h}{\lambda d(p)(\theta-\lambda)}$$
 (12)

From this equation by substituting known values $C, \lambda, h, \theta, d(p)$ and K we can find the optimum value of T .

An approximate solution to (12) can be obtained by using a truncated Taylor series expansion for exponential function as θ and λ are very small. Using Taylor series expansion equation (12) reduces to

$$T_p = \left[\frac{2K\lambda d(p)}{h\theta - (\theta - \lambda)(C\lambda + h)} \right]^{1/2}$$
 (13)

The effect of variation in perishability and price changes on the optimal order decision can be obtained from equation (9) and (13) we can get

$\frac{Q_{T_p}}{T_p} \cong \frac{1}{d(p)} \left[1 - (\theta - \lambda) \frac{T_p}{2} \right]$

The sensitivity of the order rate to change in the perishability is determined by

$\frac{\partial}{\partial \lambda} \left[Q_{T_p} / T_p \right] = \frac{\lambda}{2d(p)} + \frac{\lambda}{2d(p)} \cdot \frac{\partial T_p}{\partial \lambda} > 0$

For optimal price decision, consider the profit rate function for a fixed period length.

$f(T, p) = p \cdot \frac{e^{-\theta T}}{d(p)} - C(T, p)$

Differentiating with respect to p , we get

$$\frac{\partial f}{\partial p} = \frac{e^{-\theta T}}{\{d(p)\}^2} [d(p) - p d'(p)] + \frac{d'p}{[d(p)]^2} \left[\frac{C\{1 - e^{-(\theta-\lambda)T}\}}{T(\theta-\lambda)} + \frac{h}{T(\theta-\lambda)} \left\{ T - \frac{\theta T^2}{2} \right\} + \frac{e^{-(\theta-\lambda)T}(e^{-\lambda T} - 1)}{\lambda} \right]$$

Also $\frac{\partial f}{\partial p} = 0$ implies

$$\frac{d(p)}{d'(p)} + e^{\theta T} \left[\frac{C\{1 - e^{-(\theta-\lambda)T}\}}{T(\theta-\lambda)} + \frac{h}{T(\theta-\lambda)} \left\{ T - \frac{\theta T^2}{2} \right\} + \frac{e^{-(\theta-\lambda)T}(e^{-\lambda T} - 1)}{\lambda} \right] = p_T$$

IV. EXAMPLE AND TABLES

SPECIAL CASE

Case 1 If $\theta = 0$ and Demand = $d^*(p)$ (say)

Then this model reduces to Cohen model (2).

Case 2 If $\theta = 0, \lambda = 0$ and demand is constant. Then this model reduces to the standard formula for non decaying inventory.

NUMERICAL EXAMPLE

Consider an inventory system such that

$K = \text{Rs. } 50$ per order, $h = \text{Rs. } 0.50$ per unit per week, $d(p) = 25 - 0.5 p$.

The optimum or time period.

Table 1 Tabulation of T_p for different values of the parameter.

λ	θ	P^*	C	T_p
0.02	0.0	40	30	21.32
0.06	0.02	40	30	17.14
0.06	0.02	30	30	24.25
0.10	0.02	30	40	16.43
0.10	0.06	30	40	21.82

We have found the values of T_p for a fixed set of values

K and h varying values of C, p, θ and λ . Table 1

indicates that with the increasing λ, T_p decreases and with the increasing value of θ, T_p increases and the same will increase with the decrease in p .

V. CONCLUSION

In this paper, we developed a deterministic inventory model for deteriorating items when demand is a function of time and price both.

The result of the model is important for formulating the decisions when the inventory decay with constant rate and demand is a function of time and price both.

Two special cases illustrates the effectiveness of the developed model. The future study will incorporate any factual relation that may exists between time and price both

in the demand rate function and variable rate of deterioration.

AKNOWLEDGEMENT

We are thankful to Dr.A.P.Singh of SGRR (PG) College Dehradun for his constant support and review of our research paper. We also thank Prof. Yudhveer Singh Moudgil of Uttranchal Institute of Technology for reviewing our research paper.

REFERENCES

- [1] Cohen M.A. (1976). Analysis of single critical number ordering policies for perishable inventories. *Operations Research* **24**, 726-741.
- [2] Cohen M.A. (1977). Joint pricing and ordering policy for exponentially decaying inventory with known demand. *Nav. Res. Log. Quar.* **24(2)** 257-268.
- [3] Emmons H. (1968). A replenishment model for radio active nullide generators. *Management Science* **14**, 263-274.
- [4] Ghare P. M. and Schrader G. F. (1963). A model for an exponentially decay inventory. *The Journal of Industrial Engineering* **14**, 238-243.
- [5] Gupta P. N. and Jauhari R. (1993). Optimum ordering interval for constant decay rate of inventory, discrete-in-time. *Ganita Sandesh* **7(1)** 12-15.
- [6] Mukherjee S. P. (1987). Optimum ordering interval for time varying decay rate of inventory, *Opsearch* **24(1)** 19-24.
- [7] Naresh Kumar and Sharma A. K. (1999). On deterministic production inventory model for deteriorating items with an exponential declining demand. Presented at 11th RGP held at SGN Khalsa College, Sri Ganga Nagar (Dated 21, 22 Nov. 1999).
- [8] Naresh Kumar and Sharma A. K. (2009). Optimum ordering interval with known demand for items with variable rate of deterioration and shortages (Accepted for publication in A.S.P.).

Study of Implementing 5S Techniques in Plastic Moulding

Prof. S. B. Khedkar¹, Prof. R. D. Thakre², Prof. Y. V. Mahantare³, Mr. Ravi Gondne⁴

^{*}(Mechanical Engg Department, B. D. College of Engg Sevagram, India)

^{**}(Mechanical Engg Department, B. D. College of Engg Sevagram, India)

ABSTRACT: 5S is a basic foundation of Lean manufacturing systems. It is a tool for cleaning, sorting, organizing and providing the necessary groundwork for workplace improvement. This research effort dealt with the implementation of 5S methodology in the S.P. Plastic Industry MIDC, Hingna Road, Nagpur 16. A detailed application of the 5S system is given. It will impact the instructors and workman of Industry that work within the selected place. By following the 5S methodology, this research effort may show significant improvements to safety, productivity, efficiency, and housekeeping. The research documents improvements by using before and after pictures. It also intends to build a stronger work ethic within the workman and engineer who would be expected to continue the good practices.

Keyword- 5S, Productivity, Quality management

I. INTRODUCTION

Plastic moulding is very widely used process in today's world. Due to the high cost of metallic materials and their lack of availability the plastic materials are very widely used in many applications.

The plastic material has very good qualities and their qualities can also be improved easily by the use of catalyzers. Thus plastic moulding is getting very high attention. Plastic moulding has very wide classification. Plastic products with simple shapers are easy to manufacture. But products with complex shapes are difficult to manufacture the injection moulding is used.

Injection moulding machines are used for their high productivity and high quality of production. In injection moulding the products which require uniform thickness with high density are easy to manufacture. z The plastic moulding industries are replacing their conventional machines by the injection moulding machines due to their high productivity and user friendly control.

In S.P. Plastic Industries by conducting study that helped us to identify the problems with the machines, environments, safety and cleanness of the industry.

5S is a system to reduce waste and optimize productivity through maintaining an orderly workplace and using visual cues to achieve more consistent operational results. The term refers to five steps – sort, set in order, shine, standardize, and sustain – that are also sometimes known as the 5 pillars of a visual workplace. 5S programs are usually implemented by small teams working together to get materials closer to operations, right at workers' fingertips and organized and labeled to facilitate operations with the smallest amount of wasted time and materials.

"A place for everything, and everything in its place" is the mantra of the 5S method, and storage and workspace systems such maximum use of cubic space for the highest density storage. The result is an improved

manufacturing process and the lowest overall cost for goods produced.

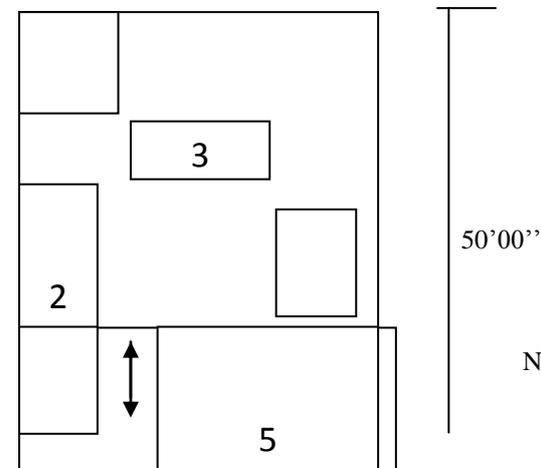
II. COMPANY INFORMATION

Name of Industry: S.P. Plastic Industry
Address: 39, Suvarna Laghu Udyog Yojana,
Electronic Zone Chowk,
MIDC, Hingna Road, Nagpur 16.

2.1 Product produces

- Mug
- Fruit Container
- Buckets
- Bowls etc

2.2 Plant layout



1. 1st floor office: Owner and other official work are done.
2. Raw material storage area.
3. Injection moulding machine 1.
4. Injection moulding machine 2.
5. Dispatch area.

2.3 Raw material

Poly Propylene (Pp) Grannules

Types: Cpp (Colour Poly Propylene), Wpp (White Poly Propylene, Marble Quality), Hd (High Density), Ep (Engineering Plastic).

2.4 Process

This is small industry. In which plastic is used as raw material and this raw material is converted by injection moulding process into the finish good or different parts. The process for the different parts is given below.

1. When raw material comes into the company, it is store in the storage by using fork left (manually).

2. After stage of material the quality inspector check the raw material visually. If any problem found in the raw material then these raw materials send back or salvage it.
3. The raw material fed into the hopper manually. According to part has to be done.
4. For the process injection moulding machines is used in which melted raw material is converting into the finish product by using die cavity.
5. The raw of injection moulding machines fed raw material to the torpedo. The torpedo heat the raw material and convert it into melted plastic, this melted plastic fed through the nozzle in the die cavity and generate the profile according to shape generate into the die cavity.
6. After completion of process the finish good removes from the die and the operator operate machines for next part.
7. The operator removes the exceed material from the finish good by using hacksaw blade.
8. This finish product carry to the storage area manually where quality check visually.
9. Then after packing is done in the storage area.

III. AREA OF STUDY

Proper utilization of storage space more prominently.

3.1 Objective of study

To improve the productivity by introducing 5S system in the work place.

3.2 Problems identification

While studying identified problems are

1. Utilization of stored space for finished product is not proper.
2. Utilization of stored space for raw material is not proper.
3. More time is required for packing of finished product.

Analysis and elaboration of the problem

3.2.1 Utilization of stored space for finished product is not proper:

As per our study it is found that the total build up area of the plant was not properly utilized. it is also found that the storage space was not enough to store the finished product. The production in one day was occupying more than the hall of the space of the storage space. Thus the total production of the day has to send to market to store the ongoing production. This space utilization of storage space can be done more efficient.

3.2.2 Utilization of stored space for raw material is not proper:

Utilization of space due to unwanted raw material and there is no specified space for raw material storage.

3.3.3 More Time Required for Packing of Finished Products: The finished product of the machine has to be well packed before it reaches the market because, the quality of the product has to be good and no damage should be there on the product. The packing of finished product was collected in a lot of thousand and then packed in big polythene bags. The bags are so large that the

labours were not able to handle the bags properly and set the goods in the bags. This process was not properly done, and was time consuming. The process could be done by the automatic packing machines

4.5 'S' SYSTEM

Establish and maintain a clean, neat and tidy workplace Translation of 5 Japanese S's, what is 5S and why do we want to do it? 5S represents 5 disciplines for maintaining a visual workplace (visual controls and information systems).

These are foundational to Kaizen (continuous improvement) and a manufacturing strategy based "Lean Manufacturing" (waste removing) concepts.

5S is one of the activities that will help ensure our company's survival.

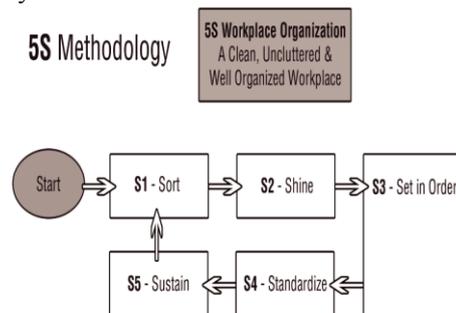


Fig.1

4.1 Sort / Arrangement (SEIRI)

(Eliminate unnecessary items)

Through the suitable sorting it can be identified the materials, tools, equipment and necessary information for realization the tasks. Sorting eliminates the waste material (raw materials and materials), nonconforming products, and damaged tools. It helps to maintain the clean workplace and improves the efficiency of searching and receiving things, shortens the time of running the operation. The 1S rule's proceedings

A) On the first stage one should answer to so-called Control Questions:

- Are unnecessary things causing the mess in the workplace?
- Are unnecessary remainders of materials thrown anywhere in the workplace?
- Do tools or remainders of materials to production lie on the floor (in the workplace)?
- Are all necessary things sorted, classified, described and possess the own place?
- Are all measuring tools properly classified and kept?

On the basis of the answer to the above questions it is possible the estimation of the workplace in terms of the 1S rule so littering the workplace. If on any question answer is yes, it should execute sorting of things, which are in the workplace.

B) On the second stage one should execute the review of all things which are in the workplace and group them according to the definite system. According to carried out sorting it should execute the elimination from the workplace the things, which were found „unnecessary”.

C) To permanent usage the 1S rule is so-called the Programmed of the Red Label. It means giving the red label to things, which operator will recognize as useless

within his workplace. This label will make possible not only the elimination of the given thing, but through its own formula will make possible the liquidation of the reasons of appearing on the workplace this given thing.

4.2 Set in Order / Neatness (SEITON)

(Efficient and effective storage method)

Especially important is visualization of the workplace (eg. painting the floor helps to identify the places of storage of each material or transport ways, drawing out the shapes of tools makes possible the quick putting aside them on the constant places, colored labels permit to identify the material, spare parts or documents etc.).

Implementing the 2S rule

It should execute the segregation of things and mark the places of their storing. Used things should always be divided on these, which should be:

- In close access (1st degree sphere),
- Accessible (2nd degree sphere),
- In the range of hand (3rd degree sphere).

To the estimation of the workplace in terms of the 2S rule that is setting in order things serve the following Control Questions:

- Is position (location) of the main passages and places of storing clearly marked?
- Are tools segregated on these to regular uses and on specialist tools?
- Are all transport palettes storage on the proper heights?
- Is anything kept in the area of devices against the fire?
- Has the floor any irregularity, cracks or causes other difficulties for the operator's movement?

Things used occasionally and seldom should be on the workplace but outside the direct using sphere. Their distance and location from the place of work should depend on the frequency of using these materials or tools. Places of storage should be marked in the manner making possible their quick identification. It can be used colored lines, signs or tool boards.

4.3 Shine / Cleanliness (SEISO)

(Thoroughly clean the workplace)

Regular cleaning permits to identify and to eliminate sources of disorder and to maintain the clean workplaces. During cleaning it is checked the cleanness of machine, workplace and floor, tightness of equipment, cleanness of lines, pipes, sources of light, current data, legibility and comprehensibility of delivered information etc. Indispensable is also taking care of and maintenance the personal tidiness of the operator.

Implementing the 3S rule

The first step of realization the 3S rule is renovation the workplace. It is assumed that „the first cleaning” forces the exact checking of usage two of the previous rules. The usage of the 3S rule relies on everyday keeping in faultless cleanness the workplace. It is executed by the operator of the given workplace. To the estimation of the workplace in terms of the 3S rule, that is cleaning the workplace, serve the following Control Questions:

- Are the oil's stains, dust or remains of metal found around the position, machine, on the floor?
- Is machine clean?
- Are lines, pipes etc. clean, will they demand repairing?
- Are pipe outlets of oils not clogged by some dirt?
- Are sources of light clean?

4.4 Standardize / Order (SEIKETSU)

(Order and control to be established for)

Worked out and implemented standards in the form of procedures and instructions permit to keep the order on the workplaces. Standards should be very communicative, clear and easy to understand. Regarding this during preparation and improving, it should be involved all participants of the process on the given workplace, it means direct workers. The group knows the best specificity of its own activities, and process of elaboration and after that, usage gives them possibility of understanding the essence and each aspect of the operation. In the aim of assuring all the easy access, obligatory standards should be found in constant and visible places.

It is assumed that standards should not be implemented only in the typical operational processes e.g. production, movement maintenance, storing, but also in the administrative processes, for example: book-keeping, customer service, human resources management, or secretariat service.

4.5 Sustain / Discipline (SHITSUKE)

(Sustain new status quo 'everything in its place')

Implementing the idea of the 5S will demand from workers the compact self-discipline connected with implementing and obeying the rules of regularity in cleaning and sorting. It leads to increasing the consciousness of staff, and decreasing the number of non-conforming products and processes, improvements in the internal communication, and through this to improvement in the human relations.

It is also important to understand the need of executing the routine inspections of usage the 5S rule. This inspection is executed by helping of so-called Check List and created on its basis the radar graph of the 5S, which serves to estimation of the workplace. The inspection of realization of the 5S rule is executed once a month by chosen team implementing the 5S rule – the control team

V. IMPLEMENTATION OF 5S TECHNIQUES

5.1 Sort

All unneeded tools, parts and supplies are removed from the area. The company layout is fixed according to process but the company does not consist of a systematic arrangement for various material handling and storage

5.2 Set in Order

A place for everything and everything is in its place.



Fig. 2

5.3 Shine

The area is cleaned as the work is performed



Fig. 3

5.4 Standardize

Cleaning and identification methods are consistently applied.



Fig. 4

5.5 Sustain

5S is a habit and is continually improved the company culture

VI. BENEFITS

1. The production rate will increase due to systematic arrangement.
2. The space utilization will increase.
3. The atmospheric conditions will improve.
4. Clean and hygienic condition is achieved.
5. It is convenient to handle and operate each and every material.
6. Moral support of the operators and workmen's are improved.
7. Storage space is increased within the same area.

VII. CONCLUSION

The advantages from implementing the 5S rules 1S:

- Process improvement by costs' reduction,
 - stock decreasing,
 - Better usage of the working area,
 - Prevention of losing tools, 2S:
 - Process improvement (increasing of effectiveness and efficiency),
 - Shortening of the time of seeking necessary things,
 - Safety improvement.
- 3 S:

- Increasing of machines' efficiency,
- Maintenance the cleanness of devices,
- Maintenance and improvement of the machines' efficiency,
- Maintenance the clean workplace, easy to check,
- Quick informing about damages (potential sources of damages),
- Improvement of the work environment,
- Elimination of the accidents' reasons, 4 S:
- Safety increasing and reduction of the industry Pollution,
- Working out the procedures defining the course of processes, 5 S:
- Increasing of the awareness and morale,
- Decreasing of mistakes quantity resulting from the inattention,
- Proceedings according to decisions,
- Improvement of the internal communication processes,
- Improvement of the interhuman relations.

REFERENCES

- [1] J. Michalska and D. Szewieczek, "The improvement of the quality management by the activity-based costing", Journal - Journal of Achievements in Materials and Manufacturing Engineering 21/1(2007) 91-94.
- [2] Mihail Aurel Titu; Constantin Oprean and Daniel Grecu, Proceeding of the international multi-conference of engineer and computer scientist 2010 Vol III, IMECS 2010, Hong Kong, March 17-19, 2010.
- [3] Maharjan, Shyam K., Implementing the 5S Methodology for a Graphic Communications Management Laboratory of University of Wisconsin-Stout May, 2011, pp 20-31
- [4] Taiichi Ohno, Toyota production system, Productivity press, 1988.
- [5] S.K. Ho, TQM an Integrated Approaching – Implementing Total Quality through Japanese 5S and ISO 9000, Kogan Page, London, 1996.

Identification, assessment and verification of current technical failure modes at automated infusion systems

Thomas Lekscha¹, Daniela Durackova²

¹(Faculty of Engineering Science, Department of Medical Technology, Jade University WOE, Germany)

²(Faculty of Engineering and Information Technology, Institute of Electronics and Photonics, Slovak University of Technology in Bratislava, Slovakia)

ABSTRACT: A basic requirement for the identification, assessment and verification of current failure modes in automated infusion systems is a scientific survey of the status quo in clinical practice. The rise in critical incidents involving automated infusion technology from 2000 to 2009 was disclosed at 63 %. This remarkable rise requires a precise analysis of the error causes. The identification and evaluation of these technical failure modes is the primary aim of this study.

Keywords: Infusion pump, Survey, Technical sources of error, Verification.

I. INTRODUCTION

Today, automated infusion systems are indispensable and routinely used in clinics for the parenteral infusion of fluid substances into the blood circulation of the human body. Infusion therapy plays an important role during the treatment of patients. Almost 90 % of all patients who undergo stationary treatment receive infusion solutions. More than 600 million infusion solutions flow into the veins of patients every year.

Infusion therapy supported by medical technology is used whenever the mechanisms in the metabolism of the human body have become unbalanced. Infusion therapy enables maintaining, correcting and substituting the necessary quantities of substances. Infusion systems are used, for example, to compensate for dehydration, normalize the electrolyte metabolism, maintain the acid-base balance and to administer drugs. Automated infusion systems allow the infusion of fluids at a certain rate, certain quantity and in different types of application.

In order to minimize risks in the application of automated infusion systems, it is imperative to investigate and analyse failure modes. The aim of this study is to identify and evaluate the frequency and type of occurrence of device failures in automated infusion systems under safety-related aspects. Primarily, the technical failure modes of infusion sets will be investigated and not the possible medical hazards for the patient.

I. CURRENT STATE OF RESEARCH

At present, there is no analysis of technical sources of errors in automated infusion systems. According to the data-recall facility, only a few comparable investigations need to be examined.

The following key facts indicate the acuteness of the planned thesis:

- 90 % of all patients, treated in stationary wards, receive infusion solutions [1]

- More than 600 million infusion solutions/year flow into the veins of patients [2]
- The announced rise in critical incidents involving automated medical devices, between 2000 and 2009 in Germany was 255 % [3]
- The announced rise in critical incidents involving automated infusion technology between 2000 and 2009 in Germany was 63 % [4]
- More than 10.000 complaints were received annually with regard to infusion pumps between 2005 and 2009 in the USA [5].

In the "Patient Security Agenda 2008", the Action Alliance Patient Security stated that there are large gaps in information for the current study situation with regard to compiling safety-related data concerned with medical products [6].

II. PROCEDURAL METHOD

The aim of this study was to observe and analyse the error frequency of infusion system equipment. In order to obtain data concerning sources of errors in respect of infusion equipment, the first step here is to define the most frequently used infusion equipments. They are generally arranged according to their application technology:

- Gravity feed infusion with infusion control
- Infusion Pumps with rotating peristaltic
- Infusion pumps with linear peristaltic
- Infusion pumps with volume chamber
- Syringe pumps.

In the inquiries the different applications techniques of the infusion pumps were not differentiated, because all defined and asked sources of errors are to be found with all pump types.

This study is only concerned with the aspect of technical errors (mechanical, electronic) on infusion devices. There is no attempt to include "accidents" which have occurred during the use of such equipment on patients. The analysis of events (injury to patients due to a technical medical device) is not a part of this study.

The sources of errors considered in this study are purely technical faults within infusion equipment itself and have no direct relation to the application of the equipment on patients.

The data with which to analyse the error frequency of infusion system equipment was gathered by means of anonymous surveys in 100 hospitals throughout Germany. The surveys to be evaluated were from the years 2006 and 2010. A standard questionnaire was drawn up provided with a clear grading scheme. This ruled out

misinterpretation by those questioned to a large extent. The used questionnaire was the same for both surveys. The implementation of an anonymous questionnaire meant that no conclusions could be drawn regarding the hospital questioned (technical service centre), the persons questioned (service personnel) and the infusion pump manufacturer.

An “anonymous” questionnaire was used “deliberately” in order to compile realistic, undistorted data. By being non-traceable, staff in service centres could analyse the device errors freely. Thus, they were instruction-free and not influenced in their evaluation of the error frequencies. This fact supported compilation of a realistic analysis of the error frequencies. The persons questioned (medical engineers at the respective service centres) were requested to answer the questions with a grade between *one* and *nine*. The grading key should reflect the error frequency regarding the respective infusion pump types; whereby the grade *one* represented a very low error frequency and *nine* a very high error frequency. The grade *zero* was considered as not assessable. The first of all defined, then asked and rated Sources of Errors were:

- Drive: Motor, gear, motion unit
- Power supply: Mains cable, IEC socket, power adapter, battery
- Software: Type-specific device software
- Keyboard: Keys, switches, pressure point, membrane keyboard
- Body: Cracks, brittleness, breakage, leakage
- Holder: Fixation to infusion stands, threaded rod
- Hose system: Accessories, type-specific transfer systems
- Handling: Operating errors by personnel
- Dirt: Labels illegible, soiled sensors, smeared display.

Using the grades (1 to 9) from the predefined grading scheme ensured that only the error frequency (the frequency with which a technical error occurred) was assessed. The number of different types of infusion pumps used in the various hospitals and the frequency with which the pump systems are used on patients was not an aspect for consideration through the questionnaire.

III. EVALUATION OF THE INQUIRIES

Of the 50 questionnaires sent for the survey in 2006, 23 were returned plausibly answered and evaluated, resulting in a response rate of 46 %. Of the 50 questionnaires sent for the survey in 2010, 24 were returned plausibly answered and evaluated, resulting in a response rate of 48 %.

The relatively high and good response rate of both inquiries, were certainly partly due to the stamped addressed envelope enclosed with the questionnaire. All the questionnaires returned, were sensibly answered and could be evaluated.

Up to now, the frequency with which the technical error sources predefined in the questionnaire occurred has been evaluated. The evaluation is presented per individual infusion pump type and as an overall comparison. The

frequency of use of the various infusion pump types in everyday work in hospitals was not taken into account in the evaluation. The primary objective of the exercise was to compile the frequency of sources of errors not the frequency of use of various pump systems.

The following graph (Fig. 1) shows a comparison of the frequency of all sources of errors, year 2006 to 2010.

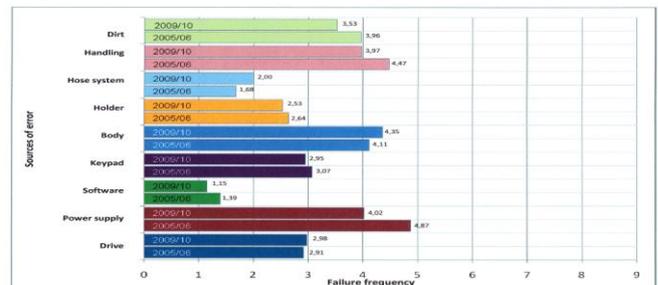


Fig. 1. Comparison, as arithmetic average, of the frequency of sources of errors of all type of pumps, survey 2006 to survey 2010

For the compiling of the statistical evaluation of the surveys, “IBM SSPS Statistics 18” software was used. Initially, the data records were gathered, samples of which were analysed with regard to normal population. The “Shapiro-Wilk test” was used here. During analysis with “SPSS”, the significances were mainly $p < 0.05$, i.e. significant and, thus, not a normal distribution.

During the subsequent group comparisons (comparison of pump types per potential source of error), an evaluation using the “Kruskal-Wallis test” was implemented for not normally distributed samples.

The significance was established as a rank comparison. The statistical evaluations have been arranged individually according to the sources of errors and totalled in the following tables. The statistical evaluation per source of error is represented by average values, standard deviations, medians and significances (p-value).

Level of significance [7]:

- $p \geq 0,05$ is equivalent to not significant
- $p < 0,05$ is equivalent to significant
- $p \leq 0,01$ is equivalent to very significant
- $p \leq 0,001$ is equivalent to very highly significant.

The following table (TABLE 1) is an example of 18 statistically evaluated tables with different sources of errors at different automated infusion systems. The evaluation of the significance of this sample table, shows a significance to "Kruskal-Willis" of $p = 0.123$. It is equivalent to not significant.

TABLE 1
Comparison of all infusion pumps, source of error “Drive”, 2006

Infusion system	Arithmetic average	Standard deviation	Median	Minimum	Maximum	N
Syringe pumps	3.04	1.22	3.00	1	5	23
Finger pumps	3.21	1.55	3.00	1	6	19
Roller pumps	2.29	0.76	2.00	1	3	7
Piston pumps	3.50	2.12	3.50	2	5	2
Infusion control	1.50	1.00	1.00	1	3	4
Total	2.91	1.37	3.00	1	6	55

In the next two tables (TABLE 2 and TABLE 3), the results of the evaluated ranking of the sources of errors, by sequence of occurrence, is shown.

TABLE 2
Weighting of the sources of errors and statistic in 2006

Weighting by rank	Survey in 2006	Statistic Test	Statistic Test
		Kruskal-Wallis	Friedman
		Significance per error source	Significance correlate
1. rank	Power supply	p= 0.017	p ≤ 0.001 very highly significant
2. rank	Handling	p= 0.569	
3. rank	Body	p= 0.007	
4. rank	Dirt	p= 0.068	
5. rank	Keypad	p= 0.089	
6. rank	Drive	p= 0.123	
7. rank	Holder	p= 0.119	
8. rank	Infusion hose system	p= 0.436	
9. rank	Software	p= 0.537	

TABLE 3
Weighting of the sources of errors and statistic in 2010

Weighting by rank	Survey in 2010	Statistic Test	Statistic Test
		Kruskal-Wallis	Friedman
		Significance per error source	Significance correlate
1. rank	Body	p= 0.269	p ≤ 0.001 very highly significant
2. rank	Power supply	p= 0.021	
3. rank	Handling	p= 0.337	
4. rank	Dirt	p= 0.458	
5. rank	Drive	p= 0.428	
6. rank	Keypad	p= 0.753	
7. rank	Holder	p= 0.149	
8. rank	Infusion hose system	p= 0.653	
9. rank	Software	p= 0.707	

The tables above give a good overview about the ranking of the sources of errors and their statistical evaluation in the different years 2006 to 2010. For analyzing the discrepancies between groups (in our case the sources of errors) the “Friedman-test” was used. The statistic test-result is shown as a very highly significance correlate of $p \leq 0.001$.

In the following tables (TABLE 4 and TABLE 5) present the contrasting juxtapositions of the sources of errors between the inquiries in 2006 and 2010. The shifting of the ranking is clearly shown by different colours.

TABLE 4
Comparative Weighting of the sources of errors in 2006 and 2010

Weighting by rank	Survey in 2006		Survey in 2010	
	Rank	Source	Rank	Source
1. rank	1	Power supply	1	Body
2. rank	2	Handling	2	Power supply
3. rank	3	Body	3	Handling
4. rank	4	Dirt	4	Dirt
5. rank	5	Keypad	5	Drive
6. rank	6	Drive	6	Keypad
7. rank	7	Holder	7	Holder
8. rank	8	Infusion hose system	8	Infusion hose system
9. rank	9	Software	9	Software

TABLE 5

Weightage of the evaluation-change of the inquiries 2006 to 2010 as ranking, positive to negative change

Weightage Evaluation-change	Change of evaluation 2006 to 2010		Sources of errors
	positive change to negative change	Δ arithmetic mean trend	
1. rank	- 0.85	↓	Power supply
2. rank	- 0.50	↓	Handling
3. rank	- 0.43	↓	Dirt
4. rank	- 0.24	↓	Software
5. rank	- 0.12	↓	Keypad
6. rank	- 0.11	↓	Holder
7. rank	+ 0.07	(↑)	Drive
8. rank	+ 0.24	↑	Body
9. rank	+ 0.32	↑	Hose system

Comparison of the evaluations of the years 2006 and 2010 indicates a clear trend within the error source rankings. Despite the distinct time gap between the two years in which the questionnaires were issued, the top three error sources (rankings) have clearly remained the most significant. Although there is a slight shift in the ranking positions (TABLE 4). The error sources “Power supply”, “Handling” and “Body” prove the most serious sources of faults in both questionnaires.

IV. CONCLUSION

The primary objectives of this study were to complete a compilation of data regarding the sources of errors which frequently occur in technical infusion devices.

The increasing use of technical infusion equipment does not only raise the demands of technical safety but also on the functionality and operational availability of the various devices. A negative consequence of the increase in use of technical devices in infusion therapy is the increased occurrence of technical error sources in everyday clinic life.

The motivation behind this study was that the stock and compilation of data on the problems outlined above, was not available and/or was very fragmentary. Research on the subject showed that the basic theme of this study is highly topical and has not been dealt with to any significant degree. All the literature sources researched indicated a definite lack of information acquired from the compilation and analysis of sources of technical errors in technical infusion equipment. Some of the relevant literature sources urgently recommended the acquisition of data to discover the gaps in error source data and fill them.

The data used in this study was gathered by means of two separate and independent, anonymous questionnaires. Both questionnaires, one from the year 2006 and the other from 2010, were completed as part surveys within a total of 100 hospitals and their service centres. The questionnaires contained standardised questions and had a fixed method of assessment. The standardised questions and defined assessment scheme was intended to eliminate any possibility of false interpretation. By issuing the identical questionnaires to different hospitals in two different years, an attempt could be made to evaluate the results. A comparison of the results shows, that an evaluation by

means of a second questionnaire was both useful and effective.

Distinct sources of errors became apparent which topped the list in a negative sense in both questionnaires. The most distinct, highest placed sources of errors were:

- Power supply
- Body (housing)
- Handling
- Dirt.

As a result of the scientific survey extracts compiled in 2006 and 2010, this study has contributed to updating the status quo in respect of error source analysis. A survey of data, of the subject described, did not exist before. The results of the questionnaires and analysis of the data have enabled the error sources related to infusion pumps to be ranked, information which can then be used for improvements in design and development.

The results of this study indicate considerable potential in the objective of minimising and preventing technical sources of errors in infusion systems by applying the relevant measures presented in this work. But the results of this study also uncover more questions in need of research which it is hoped will be answered in follow-up studies:

- 0
- Is it realistic to propose worldwide error management of error sources in associated with infusion pumps?
- On what legal foundation or platform should this error management be based?
- Can worldwide recommendations regarding the design and development of infusion pumps contribute to a reduction in the sources of errors and thus to a prevention of incidents involving patients?

These are questions, which could be answered in next studies soon.

REFERENCES

- [1] B. Braun Melsungen AG, Infusion- und Transfusions geräte, <http://www.bb Braun.de/cps/rde/xchg/bbraun-de/hs.xsl/products.html?level=1&id=00020741570002085860>, current access 27.03.2012
- [2] Alter, U., Braun Melsungen AG, *Healthcare Journal* 02/10, Page 1, Melsungen 2010
- [3] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Anzahl der Risikomeldungen 2000-2009, 53175 Bonn, Germany, http://www.bfarm.de/SharedDocs/1_Downloads/DE/Medizinprodukte/riskinfo/wissauf/statist-Auswert_quartalsweise_Anzahl-Risikomele.pdf?__blob=publicationFile, current access 15.02.2011
- [4] Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Anzahl der Risikomeldungen nach Produktgruppen 2000-2009, 53175 Bonn, Germany, http://www.bfarm.de/SharedDocs/1_Downloads/DE/Medizinprodukte/riskinfo/wissauf/statist-Auswert_quartalsweise_Anzahl-Risikomele_Produktgr.pdf?__blob=publicationFile, current access 15.02.2011
- [5] Meier, B., F.D.A Steps up Oversight of Infusion Pumps, *The New York Times*, April 2010, http://www.nytimes.com/2010/04/24/business/24pump.html?_r=1&scp=1&sq=F.D.A.%20Steps%20Up%20Oversight%20of%20Infusion%20Pumps&st=cse, current access 08.4.2011
- [6] Aktionsbündnis Patientensicherheit, *Agenda Patientensicherheit 2008 Auflage 2*, Seite 4, Institut für Patientensicherheit der Universität Bonn, November 2008
- [7] Victor, A.; Elsäßer, A.; Hommel, G.; Blettner, M., Wie bewertet man die p- Wert-Flut?: Hinweise zum Umgang mit dem multiplen Testen, *Deutsches Ärzteblatt* 4(107), Deutsches Ärzteblatt 2010

AUTHORS

Thomas Lekscha received his Dipl.-Ing. from the Faculty of Engineering Science, Jade University in Wilhelmshaven, Germany, and his M.Sc. from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava. He is currently working toward the Ph.D. degree at the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava. He is a Member of a research staff for medical technology at the Department of Medical Technology, Jade University, Germany.

Daniela Durackova received her M.Sc. from the Faculty of Engineering and Information Technology, Slovak University of Technology in Bratislava. She received her Ph.D. from the same University. Prof. Durackova is a head of VLSI group in the Institute of Electronics and Photonics. Her research activities are from biomedical applications of neural networks to VLSI design.

Applying image processing technique to detect plant diseases

Anand.H.Kulkarni¹, Ashwin Patil R. K.²

Department of Information Science and Engineering¹

Department of Computer Science and Engineering²

Gogte Institute Of Technology, Belgaum, Karnataka, India

ABSTRACT: The present work proposes a methodology for detecting plant diseases early and accurately, using diverse image processing techniques and artificial neural network (ANN).

Farmers experience great difficulties in changing from one disease control policy to another. Relying on pure naked-eye observation to detect and classify diseases can be expensive various plant diseases pose a great threat to the agricultural sector by reducing the life of the plants. the present work is aimed to develop a simple disease detection system for plant diseases. The work begins with capturing the images. Filtered and segmented using Gabor filter. Then, texture and color features are extracted from the result of segmentation and Artificial neural network (ANN) is then trained by choosing the feature values that could distinguish the healthy and diseased samples appropriately. Experimental results showed that classification performance by ANN taking feature set is better with an accuracy of 91%.

Keywords: Artificial Neural Network, Gabor Filter.

I. INTRODUCTION

Agriculture is the mother of all cultures. It has played a key role in the development of human civilization. Agricultural practices such as irrigation, crop rotation, fertilizers, and pesticides were developed long ago, but have made great strides in the past century. By the early 19th century, agricultural techniques had so improved that yield per land unit was many times that seen in the middle ages. Agricultural production system is an outcome of a complex interaction of soil, seed and agro chemicals (including fertilizers). Therefore, judicious management of all the inputs is essential for the sustainability of a complex system. The focus on enhancing the productivity, without considering the ecological impacts has resulted into environmental degradation. Without any adverse consequences, enhancement of the productivity can be done in a sustainable manner.

Plants exist everywhere we live, as well as places without us. Many of them carry significant information for the development of human society. As diseases of the plants are inevitable, detecting disease plays a major role in the field of Agriculture. Plant disease is one of the crucial causes that reduces quantity and degrades quality of the agricultural products.

Diseases and insect pests are the major problems that threaten pomegranate cultivation. These require careful diagnosis and timely handling to protect the crops from heavy losses [2]. In pomegranate plant, diseases can be found in various parts such as fruit, stem and leaves. Major diseases that affect pomegranate fruit are bacterial blight (*Xanthomonas axonopodis pv punicae*), anthracnose

(*Colletotrichum gloeosporoides*) and wilt complex (*ceratocystis fimbriata*).

Image samples of these diseases are shown in Figure 1.



Figure 1: Various diseases affecting pomegranate

Bacterial blight is the most severe disease of the pomegranate. The disease symptoms can be initially found on stem part which gradually pervades to leaves and then to fruits. On leaves, the disease starts with small, irregular, water soaked spots that are 2 to 5 mm in size with necrotic centre of pin head size. Spots are translucent against light. Later, these spots turn light to dark brown and are surrounded by prominent water soaked margins. Numerous spots may coalesce to form bigger patches. Severely infected leaves may drop off. High temperature and high relative humidity favors the disease. The disease spreads to healthy plants through wind splashed rains and in new area through infected cuttings.

In this work we will be focusing on three different diseases which are attacked on pomegranate crop.

- 1) Alternaria.
- 2) Bacterial blight.
- 3) Anthracnose.
- 4) Fruit Anthracnose.
- 5) Stem Anthracnose
- 6) Fruit Bacterial blight.

II. METHODOLOGY

The methodological analysis of the present work has been presented pictorially in Figure 1. The work commence with capturing images using cameras or scanners. These images are made to undergo pre-processing steps like filtering and segmentation. Then different texture and colour features are extracted from the processed image. Finally, the feature values are fed as input to the ANN classifier to classify the given image.

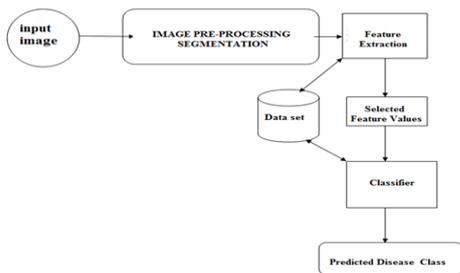


Figure 1: Block Diagram

A. Input Image: The first step in the proposed approach is to capture the sample from the digital camera and extract the features. The sample is captured from the digital camera and the features are then stored in the database.

B. Image Database: The next point in the project is creation of the image database with all the images that would be used for training and testing. The construction of an image database is clearly dependent on the application. The image database in the proposed approach consists of 140 image samples. The image database itself is responsible for the better efficiency of the classifier as it is that which decides the robustness of the algorithm.

C. Image Pre-processing: Image pre-processing is the name for operations on images at the lowest level of abstraction whose aim is an improvement of the image data that suppress undesired distortions or enhances some image features important for further processing and analysis task. It does not increase image information content. Its methods use the considerable redundancy in images. Neighbouring pixels corresponding to one real object have the same or similar brightness value. If a distorted pixel can be picked out from the image, it can be restored as an average value of neighbouring pixels. In the proposed approach image pre-processing methods are applied to the captured image which are stored in image database.

i. Segmentation

Image segmentation is process i.e. used to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze[2]. As the premise of feature extraction and pattern recognition, image segmentation is one of the fundamental approaches of digital image processing. Image Segmentation is the process that is used to distinguish object of interest from background. The proposed approach uses CIE L*a*b*, or CIELAB, color scale for use. It was intended to provide a standard, approximately uniform The CIELAB color scale is an approximately uniform color scale. In a uniform color scale, the differences between points plotted in the color space correspond to visual differences between the colors plotted. The CIELAB color space is organized in a cube form. The L* axis runs from top to bottom. The maximum for L* is 100, which represents a perfect reflecting diffuser. The minimum for L* is zero, which represents black. The a* and b* axes have no specific numerical limits. Positive a* is red. Negative a* is green. Positive b* is yellow. Negative b* is blue. Below is a figure representing the CIELAB color space.

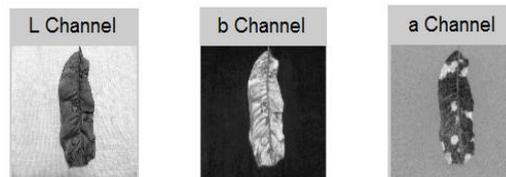


Figure 2: segmentation using CIELAB space color

D. Feature Extraction

The aim of this phase is to find and extract features that can be used to determine the meaning of a given sample. In image processing, image features usually include color, shape and texture features [3].

The proposed approach considers Gabor filter to calculate feature sets

• *Gabor filter*

A set of features are computed from the response of the image samples to the Gabor filters. They are unichannel features given by

$$e_{imn} = \sqrt{\left(\sum_{x,y} h_{imn}^2(x,y) \right)}$$

where 'e' is the energy in the filtered image. The interchannel features between different spectral channels i and j with m and m' denoting the scales of the filters is computed as

$$o_{ijmm'n}^2 = 2 - 2 \underbrace{\sum_{x,y} \frac{h_{imn}(x,y)h_{jm'n}(x,y)}{e_{imn}e_{jm'n}}}_{C_{ijmm'n}}$$

where $C_{ijmm'n}$ is the zero offset normalized crosscorrelation between $h_{imn}(x,y)$ and $h_{jm'n}(x,y)$.

D. Recognition & Classification:

The recognition process consists of two phases, training and classification. Classification of image is done ANN (Artificial Neural Network)

i. Artificial Neural Network

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process the information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze [4].

III. EXPERIMENTAL ANALYSIS & RESULTS

EXPIREMENTAL ANALYSIS

- *Experimental Analysis WRT Number of Hidden Neurons v/s Neural Network efficiency.*

Number of Hidden Neurons	Recognition Rate for. <u>Alterneria</u> (%)	Recognition Rate for BBD (%)	Recognition Rate for <u>Anthractnose</u> . (%)	Overall Neural Network Efficiency (%)
10	78	72	81	77
20	77	83.5	96	85.5
30	72.5	82	94.5	83
40	78	86	94	86
50	81.5	94	97.5	91

Table 1: Number of hidden Neurons v/s NN efficiency.

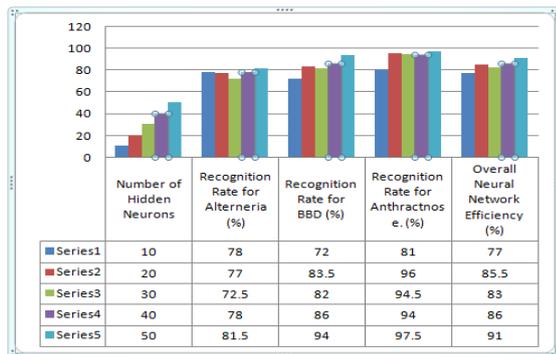


Figure 3: Graphical Analysis for Hidden Neurons v/s NN efficiency.

The table 1 shows the dependency of the efficiency on the number of hidden layers. Number of hidden layer represents number of states of the neurons in the network. The efficiency of the network is optimum when there are at least n x n numbers of hidden layers. The n here represents number of features per training set. The figure 3 shows the graphical representation of analysis with respect to Number of Hidden Neurons v/s Neural Network Efficiency which shows the network is optimum when 50 hidden neurons are considered.

- *Experimental Analysis WRT Termination error rate v/s Neural Network efficiency*

Termination Error rate (ms)	Recognition Rate for. <u>Alterneria</u> (%)	Recognition Rate for BBD (%)	Recognition Rate for <u>Anthractnose</u> . (%)	Overall Neural Network Efficiency (%)
0.1	78.5	73	81.5	79
0.01	76.5	83.5	96	85
0.001	73	82.5	94.5	84
0.0001	77	86	93	84
0.00001	81.5	94	97.5	91

Table 2: Termination error rate v/s Neural Network efficiency

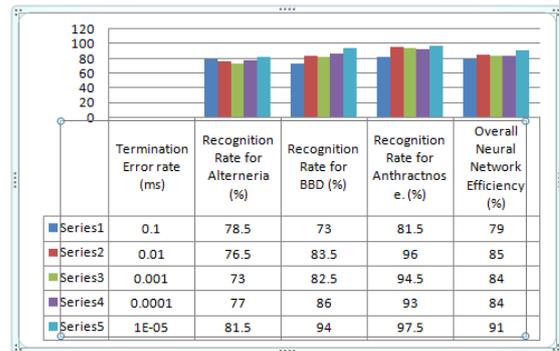


Figure 4: Graphical Analysis for Termination error rate v/s ANN efficiency

The table 2 shows the dependency of the efficiency on the termination error rate. Termination error rate represents the maximum tolerable error in classifying the values in a neural network. The efficiency of the network is optimum for more termination rate, better is the performance of the neural network. The figure 4 shows the graphical representation of analysis with respect to Termination Error Rate v/s Neural Network Efficiency which shows the network is optimum when termination error is set to 0.00001.

IV. RESULTS

In this approach, the network is trained on 140 samples from which 8 samples are alterneria, 26 samples are BBD and 89 samples are Anthractnose are used for training and testing. The Below table 3 shows the recognition rate for diseases by setting the parameters as specified in table 4.

Diseases	Recognized Samples	Misclassified samples	Recognition Rate (%)
<u>Alterneria</u>	6	2	75
BBD	21	5	80.76
<u>Anthractnose</u>	86	3	96

Table 3: Recognition rate of the diseases with uniform background

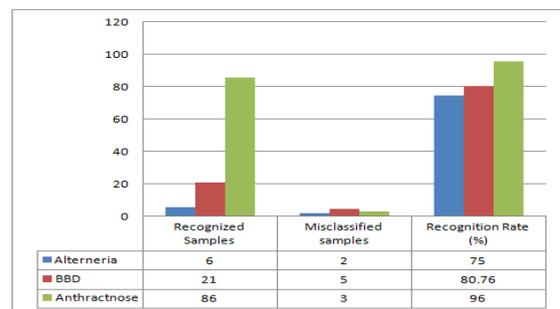


Figure 5.2.1: Recognition rate of the diseases with uniform background

Structure of Neural Network for proposed approach

Number of Input Neurons	168
Number of output Neurons	4
Hidden Layers	50
Iterations	4500
Termination Error Rate (ms)	.00001
Gradient	0.02
Recognition Rate of Neural Network (%)	91

Table 4: Neural Network Design

From the above conducted analysis the table 4 clearly shows that the performance of neural network is not only depending upon the number features; number of hidden neurons and the termination error rate but also depends on the quality of sample image. Hence an optimization must be tested with number of feature values, number of hidden neurons and the termination error rate in various different input conditions in order to correctly classify samples to their corresponding classes. The system performance can be evaluated based on its ability to correctly classify samples to their corresponding classes. Hence the above shown experimental analysis shows that the efficiency of the network is better when number of features are 168 for an image, number of hidden neurons are 50, termination error rate is .00001 and images with uniform background in light environment with minimum distance of 1 or 2 feet between the input image and the camera.

V. CONCLUSION

In this project work the area of plant diseases recognition is introduced. The system developed here is for plant diseases recognition, the development of good classification methods and precise features is very important in order to run the system in real time. Therefore proposed approach which is based on Gabor filter for feature extraction and ANN classifier for classification got a better results and recognition rate up to 91%.

An ANN based classifier is adopted which uses the combination of color and texture features to recognize and classify different plant diseases. The results are encouraging and promise the development of a good machine vision system in the area of recognition and classification of plant diseases. The proposed approach can significantly support in recognizing normal and affected produce.

REFERENCES

- [1] Dheeb Al Bashish, Malik Braik and Sulieman Bani-Ahmad, "Detection and Classification of Leaf Diseases using K-means-based Segmentation and Neural-Networks-based Classification", *Information Technology Journal* 10(2): 267-275, 2011.
- [2] Rafael C.Gonzalez, Richard E.Woods, Steven L.Eddins, *Digital Image Processing*, Pearson Education, 2nd Edition
- [3] Image Processing Toolbox™ 7 User's Guide, ©Copyright 1993–2010 by *The MathWorks, Inc.*
- [4] Christos Stergiou and Dimitrios Siganos, "Neural Networks". http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.
- [5] Gonzaliz RC., Woods RE., "Digital image processing", Pearson Education Inc., 2002
- [6] Integrating Diagnostic Expert System with Image Processing via Loosely Coupled Technique Mohammed El-Helly Central Lab. For Agricultural Expert System (CLAES) Agriculture Research Center (ARC)
- [7] El-Helly M., Onsi H., Rafea A., El-Gamal S., "Segmentation Technique for Detecting Leaf Spots in Cucumber Crop Using Fuzzy Clustering Algorithm", 11th International Conference on AI (ICAIA), 2003.
- [8] H.Al-Hiary, S.Bani-Ahmad, M.Reyalat, M.Braik & Z. Al Rahamneh [2011] Fast & accurate detection & classification of plant diseases, international journal of computer applications(0975-8887), volume 17- no.1, pp-31-38.
- [9] Advances in image processing for detection of plant diseases jayamala k. patil , raj kumar.
- [10] Color Image Classification Using Gabor Filters Vidya Manian and Marcel J. Castro Advisor: Ramón Vásquez Computing Research Conference May 4, 2000
- [11] Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction Basvaraj .S. Anami1, J.D. Pujari2, Rajesh.Yakkundimath31, 3K.L.E.Institute of Technology, Hubli – 580 030 INDIA2S.D.M.College of Engg. &Tech, Dharwar – 580 008, INDIA
- [12] http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm
- [13] Anil Kumar Singh, "Precision Farming", Water Technology Centre, I.A.R.I, New Delhi.
- [14] R. Gonzalez and R. Woods, "Digital image processing", 3rd Edition 2008, Prentice Hall, New York.
- [15] Camargo, J.S.Smith, "An image-processing based algorithm to automatically identify plant disease visual symptoms", *Biosystems Engineering* 102 (2009), 9-21.

Water Resources Management Options for Climate Change Adaptation toward Sustainable Development in Nigeria

Oloruntade, A. J. and Afuye, G. G.

Department of Agricultural Engineering Technology, Rufus Giwa Polytechnic, Owo, Ondo State, Nigeria

ABSTRACT: *The present causes and impacts of climate change on the socio-economic development of Nigeria are examined. Apart from the natural causes of climate change, many anthropogenic factors which included clean clearing, bush burning, combustion of fuels, gas flaring, etc. that are capable of releasing greenhouse gases (GHG) were highlighted. Extreme events such as floods and droughts attributable to climate change were found to have caused serious water resources problem culminating in economic loss in the country. Nevertheless, a thorough appraisal towards adaptation to climate change showed little commitment, improperly coordinated policies and poorly funded institutions which are highly inefficient coupled with general poor enforcement of relevant environmental laws. Poor knowledge of the critical importance of water resources management to climate change adaptation in the country has also been identified as a major cause of poor adaptive measures by many stakeholders. Hence, it was recommended that appropriate laws that seek to protect all water bodies should be enacted and enforced to prevent water pollution. Also, investment in the construction of adequate water storage facilities, improved capacity and information services by all relevant government agencies, conservation tillage practice and mainstreaming of adaptation to climate change through water resources management are critical to achieving sustainable development in Nigeria.*

Keywords: GHG, Floods, Droughts, Water Storage, Conservation Tillage

I. INTRODUCTION

No doubt, many Nigerians are still skeptical on whether the word 'climate change' is relevant to the economy of the country. This is because, many believe that since Nigeria and indeed Africa is less industrialized, there is minimal release of greenhouse gases (GHG) to the atmosphere which can invariably deplete the ozone layer and thus lead to global warming and climate change. However, climate change has become the reality of the moment; a consensus exists amongst scientists that global climate is changing and that if the current trends of global warming continues, rising temperature, sea level rises, and more frequent extreme weather conditions (floods, droughts, storms, heat-waves, etc.), would be witnessed in many parts of the world [1]. IPCC in [2], Fourth Assessment Report (AR4) gave the most acceptable definition of climate change as a change in the state of the climate that can be identified (e.g. by using statistical tests) by changes in the mean and/or the variability of its properties, and that persists for an extended period typically decades or longer. The present regional changes in the frequency and intensity of extreme weather events according to [3] are as a result of the climate change impact on global water resources. Higher temperatures

and changes in extreme weather conditions can affect availability and distribution of rainfall, snowmelt, river flows and groundwater, and further deteriorate water quality.

One of the most vulnerable regions in the world to climate change is Africa [2, 4], perhaps due to its natural legacy of extreme high rainfall variability [5]. The impacts of climate change is expected to be more felt in Nigeria through sea level rise along her 800 km long coast time, worsening desertification, erosion and flooding disasters and general land degradation [6, 7]. High vulnerability of the country to climate change might have arisen from over dependence on rain-fed agriculture, compounded by factors such as widespread poverty, unemployment and weak capacity. Thus, many studies have revealed that the prevailing ecological problems being faced by the country are as a result of climate change [8, 9]. Of course, recent flood events in Lagos, Oyo, Akwa-Ibom and Plateau states, to mention but a few have testified to the various scientific predictions that more negative impacts of climate change would be experienced by developing countries through water [4, 10]. In addition, many researchers have also predicted more setbacks for the country in the nearest future as a result of its seeming lack of the financial capacity and technological know-how to combat the projected negative impacts of climate change [7, 11]. If the situation is allowed to continue, the quest of the country to achieve all-round development in line with the Millennium Development Goals and Vision 20:20:20 will be greatly hindered. Hence, climate change adaptation through water resources management for sustainable development becomes an important goal that must be achieved, since water itself is part of the problem and part of the solution [10].

Nearly all aspects of the Nigerian economy, in particular health, food production and security; domestic water supply and sanitation; energy and industry; and environmental sustainability, can be greatly influenced by water resources management. Consequently, management of water resources should be properly embraced to enhance poverty reduction targets and sustainable development in all economic, social and environmental dimensions. Moreover, adaptation which refers to all the responses to climate change that may be used to reduce vulnerability [12] has been generally recommended for the developing nations of the world. However, there is close link between adaptation to climate change and water, both of which play important role in sustainable development. Furthermore, since many economic activities in Nigeria thrive on water resources, the recognition of the critical role of water resources management in adaptation and respond accordingly therefore provides opportunities for development. This is because adaptation measures that deal with climate variability which rely on the existing land and water management practices have the potential to create resilience to climate change and to improve water security and thus

directly contribute to development. Despite the proven effectiveness of these measures, they have not been adequately embraced by stakeholders, hence the need for proper re-orientation and education. Thus, this paper is aimed at examining the present causes and impacts of climate change in Nigeria, assess the various adaptation measures through water resources management and make recommendations with a view to achieving sustainable development in the country.

II. CAUSES OF CLIMATE CHANGE IN NIGERIA

Climate may change naturally, but the recent changes in climate in many parts of the world are as a result of anthropogenic factors. Fasona and Omojola in [13] stated that many adverse climatic and environmental impacts that occur today have been attributed to man's careless modifications to climate on local and to a limited extent, regional scale in some activities of the distant past. Most disasters (including flood, droughts, desertification, land degradation, subsidence, etc.) are not random events without underlying causes; they are sudden manifestations of gradual degradation processes [14]. Recent rise in population growth in Nigeria has increased the level of consumption of resources, technological advancement/changes technological advancement/changes as well as changes in organization of human societies. This has further led to changes in land use and land cover (urbanization, agriculture and forestry), biodiversity loss, changes in the composition of the atmosphere and climate change. Expansion of town and cities has brought about the need to clear many hitherto forested areas, thus reducing forest trees and the ability of the environment to absorb greenhouse gases (GHG) such as carbon dioxide. On the other way round, agricultural practices in Nigeria is synonymous with slash and burn method, especially amongst the peasants. The practice releases the dreaded GHG to the atmosphere and as well leads to desertification in less forested sudano-sahelian zone of the country.

Besides, increased concentration of GHG in the atmosphere through the industrial activities of many developed and industrialized nations has caused global warming that is being experienced across the world. Prominent amongst the GHG are carbon dioxide, (CO₂), methane (CH₄), Chlorofluorocarbons (CFC) and Nitrous oxides (NO_x). This is one of the reasons why the stance of China, a leading emitter of greenhouse gases in the world, was said to be crucial in efforts to create a successor to the current Kyoto Protocol which expires in 2012 [15]. At present, as earlier mentioned, Nigeria also contributes to the GHG emission in her own small ways. Apart from bush burning, combustion of organic fuels such as coal, oil, gas, etc. are also sources of GHG. It could be recalled that there has been increased usage of these fuels in recent time in the country owing to increased number of vehicles, use of generators and rapid industrialization. In fact, it has been revealed by many researchers that anthropogenic factors such as urbanization, deforestation, population explosion, industrialization and the emission of GHG contribute to the depletion of the ozone layer and its associated global warming and climate change [16-18].

The pursuance of food security which calls for increased agricultural activities in the country has also

necessitated increased use of nitrogen fertilizers and other agrochemicals that are capable of contributing to GHG emission. Clean clearing that is also common amongst many farmers in Nigeria for agricultural production, turns forested area to croplands thereby reducing the rate of carbon sequestration (trapping and absorption of carbon (IV) oxide gas). Thus, agriculture has therefore been described as the greatest culprit of climate change causing significant effects through the production and release of GHG and reduction of carbon sequestration [19]. Furthermore, livestock management, polythene (nylon) manufacturing, burning of organic matter and animal dung, all produce methane gas. In addition, many hydrocarbons which constitute about 14% GHG are produced for use in both domestic and the industrial sector, while gas flaring a common practice in the country has also been recognized as a major source of GHG emission in Nigeria [6], producing about 2.5 billion cubic standard feet, which is about 80% of produced associated gas in about 150 flaring sites in the Niger Delta [20]. This practice has further worsened the country rank in the world as a major emitter of GHG [6]. Indeed, Nigeria contributes more greenhouse gases than all other sources in sub-Saharan Africa combined [20, 21] mostly through gas flaring at the rate that can easily be referred to as the worst around the world [6].

III. IMPACTS OF CLIMATE CHANGE ON WATER AND ECONOMY IN NIGERIA

Water is the medium through which many of the anticipated impacts of climate change will operate [10]. Reports have shown that extreme events such as floods and droughts, which are capable of negative impacts on lives, livelihoods, land values and investment incentives in vulnerable areas, have increased in frequency and intensity in many parts of Nigeria over the past few years [12]. The River Niger Inland Delta has reduced in size from 37,000 km² in the early 50s to about 15,000 km² since 1990 and Lake Chad which earlier covered an area of 20,000 km² before 1970 now measures less than 7,000 km² [22]. This has led to the degradation of water quality and modification of the life cycles of aquatic animals; there is also the depletion of fish stocks from about 100,000 tons per annum in the 70s to less than 60,000 tons today which is likely to reduce further in the future [20].

In addition, there is great threat to food security, poverty eradication and livelihood in the country as a result of climate change as agriculture which provides job for about 60% of the citizens has become vulnerable. Many human activities that are related to natural environment like fishing, grazing lands and forests which form the source of income for many households for their basic needs have gone into jeopardy. While flooding has hindered the ability of many fishermen to fish, drought and desertification are albatross to grazing for many pastoral farmers. This has further confirmed scientific reports that poor people are more negatively affected by climate change [2, 11]. In terms of GDP, climate change is expected to cost Nigeria some fortunes, as the country's loss is projected to be between 2 and 11% of Nigeria's GDP by 2020, soaring to between 6 and 30% by the year 2050 [23]. In the area of human health, [24] has reported the ability of floods to cause potential increase in the transmission of several

water-borne and vector-borne diseases. Furthermore, global warming which is synonymous with climate change is providing a favourable environment for the breeding of vectors of water-borne diseases like malaria and cholera which can have negative implications for the health of the citizen and further reduce their productivity. This was the situation in Ibadan in 2010 when cholera epidemic was widely recorded just after the flood that ravaged the city. A vicious circle of poverty can then arise from this, since according to common saying, health is wealth. Invasion of some inland waterways in the country by some alien species like water hyacinth, water lettuce and typha which are detrimental to navigation, fishing and functional irrigation system are also traceable to climate change [20].

IV. THE CURRENT WATER RESOURCES MANAGEMENT EFFORTS IN NIGERIA

Recognizing the importance of water to the issue at stake, attempts have been made in the past and are still being made to ensure efficient water resources management in the country, however many of the efforts are with little or no adequate knowledge/inclusion of anticipated climate change. Thus, many dams, reservoirs and water conveyance channels/structures are never constructed to accommodate the present hydrologic variability caused by climate change. Expectedly, incessant needs for rehabilitation, expansion and sometimes failure of such structures (as witnessed in Sokoto, Kebbi, and Taraba States, 2010) and also being recorded in Oyo State (Ibadan) since the past three years, has been the order of the day. There is frequent overflow of many rivers, especially in the northern parts of the country, over their dam thereby leading to widespread flooding of cropped lands causing incalculable economic loss.

Many institutions that have the responsibilities of managing river basins are not doing enough to arrest the situation. At present, there are 12 of such river basin development authorities in the country. By their operational techniques, their ability to cope with the present challenge is in doubt as many of them do not possess the necessary hydrological data to effectively plan for any adaptation to the impacts of climate change on the water resources of their respective basins. Apart from the various river basin authorities, the federal government has also set up the Nigerian Integrated Water Resources Management Commission (NIWARMC), there is a Special Climate Change Unit at the Federal Ministry of Environment and the Presidential Implementation Committee on the Clean Development domiciled in the Presidency in Abuja. It is unlikely that the citizens are fully aware of the activities of many of the agencies. A major problem with all these agencies is the poor capacity building and inadequate funding occasioned by unpreparedness of government to tackle the problem of water resources management and climate change in the country headlong. Similarly, the country enjoys the membership of some Trans-boundary river basin development agencies like Lake Chad Basin Development Commission, Niger Basin Authority, etc. However, that the international agencies have actually performed better is also doubtful. Furthermore, in view of the moribund nature of many of the meteorological stations in the country, dearth of required climatic and hydrological data is a major hindrance to many scientific researches in hydrology, water management, etc.

In the area of groundwater management, available evidences are to the effect that adequate balance between discharge and recharge are not being maintained. Apart from lack of adequate data on groundwater resources of the country, there is no deliberate effort to ensure its safety and quality. There has been indiscriminate and widespread sinking of boreholes especially in the present political era when every community seeks for the provision of borehole from their elected and unelected representatives as a way of alleviating water scarcity problems. In many developed and developing economies, the practice is not allowed because of the awareness that it is capable of causing imbalance in the groundwater hydrology of a whole country. Additionally, many water storage facilities like dams and reservoirs are under-utilized; storing water that is not being used efficiently. In many urban towns and cities in Nigeria where pipe borne water is provided, leakages in pipe networks is a common source of waste in addition to deliberate wastage by many consumers because water is taken from taps without paying a dime in many of the towns and cities.

An unfortunate testimony of poor water resources management in Nigeria is high rate of water pollution. Arguably, as a result of lack of awareness, many river bodies suffer perennial and deliberate pollution from both stakeholders and non- stakeholders. While the release of industrial wastes to water bodies by some industries has become a common occurrence, many households also find rivers and streams as appropriate waste disposal sites. There is also heavy dumping of animal wastes from poultry houses and other related activities on many water bodies. Oil spillage, a common phenomenon in the Niger-Delta area has caused a lot of pollution problems, thus making water unfit for consumption and leading to the loss of many aquatic lives. It is sufficient to add here that, the situation has remained so because of the apparent negligence of governments at all levels in many instances, when the citizens commit such a crime against the nature. But for Nigeria to achieve her developmental agenda as stated in both the Millennium Development Goals and the Vision 20:20:20, efforts must be geared towards combating the menace of climate change. This calls for proper re-orientation and thus throws up challenges for all professionals especially water resources engineers.

V. CONCLUSION

From the foregoing, it has become apparent that nothing much has been done in Nigeria in the area of climate change adaptation generally and water resources management in particular. Many of the policies and programmes of governments, both at the national and local levels are yet to be properly coordinated thereby making mainstreaming of climate change adaptation difficult. Duplication of agencies and roles which breeds inefficiency in the system suggests lack of policy direction. It is not likely that the country has actually identified the key areas that require immediate attention in climate change adaptation strategies. At individual levels, there is still growing skepticism as regards whether climate change is real in Nigeria and as such many cannot see the need for attitudinal change. Inadequate awareness amongst the relevant stakeholders is also a major issue that requires urgent attention.

For the country to achieve its Millennium Development Goals and Vision 20:20:20 as highlighted in its development plan, adaptation to climate change must be vigorously pursued. As earlier posited, such an adaptation policy should take note of the critical role of water, since climate change is expected to have most negative impacts on the developing world through water. Therefore, the following recommendations are made:

- (1) There is the need to enact and enforce laws that seek to protect all water bodies against refuse dumps which sometimes encourages flooding when water cannot move freely along constructed channels; all environmental laws should be revived and enforced. The people should be enlightened to treat water as an important and invaluable resource that is exhaustible like oil.
- (2) Investment in water storage and flood control facilities to improve water management should be vigorously pursued and where dams and reservoirs are built, they should be optimally used. The people should also be encouraged to embark on rain water harvesting for both domestic and agricultural uses.
- (3) Proper coordination of government agencies saddled with the responsibilities of water resources management in the country should be undertaken. Agencies like Nigeria Meteorological Agency (NIMET) should be strengthened to perform their duties beyond forecasting alone and further provide early time warning systems that will enable the citizens prepare for emergency situations.
- (4) Farmers who are perhaps the most affected group by climate change impacts should be properly educated to adopt modern farming techniques which include the use of drought resistant varieties of crops. Similarly, areas prone to seasonal floods should be avoided while also embracing conservation tillage as against the old practice of slash and burn.
- (5) Mainstreaming adaptation to climate change through water resources management is an important option which should also be urgently pursued to achieve substantial success.

REFERENCES

- [1] Christensen, J. H., Hewitson, B., Busuioic, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W-T, Laprise, R., Magan`a Rueda, V., Mearns, L., Mene`ndez, C. G., Ra`isa`nen, J., Rinke, A., Sarr, A. and Whetton, P., "Regional climate projections", In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (eds), *Climate Change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Cambridge University Press, Cambridge, UK, pp 847-940, 2007.*
- [2] Intergovernmental Panel on Climate Change (IPCC), "Fourth Assessment Report: Climate Change 2007", Geneva, Switzerland **2007.**
- [3] Huq, S. and Reid, H., "Climate Change and Development: Consultation on key Researchable issues", Climate Change Group International Institute for Environment and Development 3 Endsleigh Street, London WC1H 0DD, UK.**2005.**
- [4] IPCC, "Climate Change 2001: Impacts, Adaptation and Vulnerability, IPCC Working Group II, Third Assessment Report" , (McCarthy, J. J., Canziani, O. F., Leary, N. A., Dokken, D. J. and White, K.S. (eds), Cambridge University Press, **2001.**
- [5] Grey, D., "Water Resources and Poverty in Africa: Breaking the Vicious Circle", Inaugural Meeting of African Ministerial Council on Water (AMCOW), Abuja Nigeria, 30th April, **2002.**
- [6] Enete, I. C. and Ezenwanji, E., "Implications of climate variability on water resources of Nigeria: A review", *Journal of Geography and Regional Planning*, 4 (13):678-682, **2011.**
- [7] Ologunorisa, T. M., "In search of Climate Justice and Equity", 1st Inaugural Lecture, Osun State University, Osogbo, Nigeria, 19th January, **2011.**
- [8] Odjugo, P. A. O., "An analysis of rainfall pattern in Nigeria", *Global J. Environ. Sci.*, 4(2): 139-145, **2005.**
- [9] Adefolalu, D. O., Pam, J. L. and Habbi, H. M., "Climate change and safety of air transportation – A Nigerian perspective", *Proceedings of the International Conference on the impacts of extremeweather and climate on socio-economic development in Africa held at the Federal University of Technology, Akure, Nigeria, 11-15 Nov. 2007. pp. 1-15, 2007.*
- [10] Sadoff, C.W. and Muller, M., "Better Water Resources Management – Greater Resilience Today, More Effective Adaptation Tomorrow", A Perspective Paper Contributed by the Global Water Partnership (GWP) through Its Technical Committee, **2008.**
- [11] Odjugo, P. A. O., "Regional evidence of climate change in Nigeria", *Journal of Geography and Regional Planning Vol.* 3(6), pp. 142-15, **2010.**
- [12] Oladipo, E., "Towards Enhancing the Adaptive Capacity of Nigeria: a Review of the Country's State of Preparedness for Climate Change Adaptation", A Report Submitted to Heinrich Böll Foundation Nigeria, September, **2010.**
- [13] Fasona, M. J. and Omojola, A. S., "Climate Change, Human Security and Communal Clashes in Nigeria, Human Security and Climate Change", An International Workshop, Holmen Fjord Hotel, Asker, near Oslo, 21-23 June, **2005.**
- [14] UNEP - GRID-Arendal Environment Times No 3, GRID-Arendal, Logum Park, Service Box 706, Arendal, In: Fasona, M. J. and Omojola, A. S., "Climate Change, Human Security and Communal Clashes in Nigeria", Human Security and Climate Change, An International Workshop, Holmen Fjord Hotel, Asker, near Oslo, 21-23 June, **2005.**
- [15] The Nigerian Tribune, "China looking to bigger climate change steps", The Nigerian Tribune, 25th November, **2010.**
- [16] Buba, A. D., "Climate change and water problems in Chad Republic", *J. Arid Environ.* 3(2): 42-50, **2004.**
- [17] DeWeerd, S., "Climate change coming home: Global warming effect on populations", *World Watch*, 2(3) 6-13, **2007.**
- [18] Odjugo, P. A. O., "The impact of climate change on water resources; global and regional analysis", *Indonesian J. Geogr.*, 39(1): 23-41, **2007.**
- [19] Ozor, N. and Nnaji, C., "The role of extension in agricultural adaptation to climate change in Enugu State, Nigeria", *Journal of Agricultural Extension and Rural Development Vol.* 3(3), pp. 42-50, **2011.**
- [20] Omenai, J., "Climate Change? What is Climate Change?", A Paper Presented at a Public Lecture Organized by the Akoko Development Group, 20-21 October, **2009.**
- [21] Ogbo, A. I and Onyedinma, A. C., "Climate Change Adaptation in Nigeria: Problems and Prospects", *Sacha Journal of Environmental Studies*, 2 (1): 130-145, **2012.**
- [22] Food Agricultural Organisation (FAO), "Adaptive Water Management in the Lake Chad Basin: Addressing current challenges and adapting to future needs", *World Water Week, Stockholm*, 16-22, August, **2009.**
- [23] The Guardian, In: Enete, I. C. and Ezenwanji, E. E., "Implications of climate variability on water resources of Nigeria: A review", *Journal of Geography and Regional Planning*, 4 (13):678-682, **2011.**
- [24] WHO Fact Sheet, "Flooding and Communicable diseases; Risk assessment and preventive measures", *World Health Organisation, Communicable Disease Working Group on Emergencies, HQ.***2010.**

Characterization and Generation of Municipal Solid Waste in North Central Nigeria

¹M. B. Oumarou, ²M. Dauda, A. T. Abdulrahim¹, ³A. B. Abubakar

¹. Mechanical Engineering Department, University of Maiduguri, Borno State, Nigeria.

². National Agency for Science and Engineering Infrastructure (NASENI), Abuja, Nigeria

³. Department of Mechanical Engineering, Ramat Polytechnic, PMB: 1070, Maiduguri; Borno State, Nigeria

Abstract: The study focused on the characterization and generation of municipal solid waste in the north central part of Nigeria. Daily samples were collected and interpreted using Microsoft Excel for quantification purposes while the characterization samples were collected during the months of February, March and April and during the raining season in August. The refuse physical characteristics were then evaluated by sifting through the waste and separate it into its various physical major components. They were analyzed for proximate and ultimate composition using ASTM standards. Average moisture contents were found to vary from 26.743 to 32.593. Wood, paper, plastic and leaves were found in varying proportions and an average waste generation of 1.23 kg per person per day was found. Energy recovery is possible with an addition of supplementary fuel as sugar cane straw, weeds. The knowledge of the refuse CV alone is not enough to conclude whether the refuse will burn or not. Its moisture content is also an essential parameter.

Keywords: municipal solid waste, generation, characterization, north central Nigeria, ultimate analysis, proximate analysis, calorific value, semi-arid region.

I. INTRODUCTION

In developing countries where waste management systems are insufficient and inefficient, coupled with the expanding urban population [1, 2], the problem of refuse disposal is reaching proportions that are causes for concern. Also, the operation and management of municipal solid waste (MSW) collection services are fairly rudimentary. This is reflected in the lack of information about the quantities and types of MSW collected, the amounts recovered, recycled and/or reused and the siting of MSW disposal sites [3].

Semi-arid regions of the world are defined as the transition zones between arid and sub humid belts and precipitation is less than potential evaporation, characterized by high temperatures (30°C- 45°C) in the hottest months [4]. In Africa, these regions are mainly characterized by low annual rainfall and the whole area is mainly desertique. Added to this, is the almost inexistent and erratic power supply from the national grid. Wastes are an important source of energy presently used in the generation of electricity and at the same time making the environment clean. The National energy supplies are at present almost entirely dependent on fossil fuels and firewood which are depleting fast [5].

Diso *et al.* [6] characterized the refuse in Kano and found that 13 megawatts of electricity could be produced from wastes in 1995 covering at least the entire Kano metropolitan area. Mengiseny and Josia [7] are of the opinion too that the main problem facing policy makers in the waste management sector is how to predict the amount of solid wastes generated in the near future in order to devise the appropriate treatment or disposal mechanism. Agunwamba *et al.* [8] analyzed the waste in Onitsha (Nigeria). Also in the southern part of the country, Adeyemi *et al.* [9] investigated the impact of waste scavenger in case study of Ilorin. Jekayinfa and Omisakin [10] considered ten agricultural wastes in Nigeria to determine their energy content using the method of Association of official Analytical Chemists. Results of their analysis showed that the mean higher heating values of the wastes samples were 16505kJ/kg, 19597kJ/kg, 20647kJ/kg, 15891kJ/kg, 17303kJ/kg, 19458kJ/kg, 28203kJ/kg, 19299kJ/kg, 21392kJ/kg and 21143kJ/kg for groundnut shell, yam peels, coconut shell, mango peels, palm oil mill effluent, corn cob, cherry, orange peels, melon shell and black walnut hull respectively. All the waste samples considered have heat values greater than some well-known biomass-fuels and fall within the limit for the production of steam in electricity generation.

Paolo [11] investigated the effect of separate collection of municipal solid waste on the calorific value of the residual waste. He emphasised that separate collection plays an irreplaceable role in solid waste management and incineration. Considering the average Italian municipal solid waste composition, he proposed separate collection scenarios different from those tested; he proposed a regression model, calibrated it and teste it partially.

Ogaji *et al.* [12] investigated the municipal solid waste (MSW) generated in Port Harcourt. They found it in large quantities, but some remains as litter in parts of the municipality. Refuse is mostly buried, but some reckless open-burning ensues, so posing environmental hazards.

Due to the heterogeneous nature of MSW, it is very difficult if not almost impossible to make projections. A comprehensive characterization of municipal solid waste is crucial to the long-term efficient and economical planning for solid waste management. Often, the problem is more pronounced when the choice of the best treatment option, to dispose of the MSW, is at hand. This paper focuses on the generation and characterisation of MSW in the north central semiarid region of Nigeria within the 2006-2011 period, for the purpose of choosing the best and appropriate option to treat and dispose of these wastes. The determination of the wastes combustion characteristics and evaluation of the amount of energy to be obtained from such wastes is carried out.

II. MATERIALS AND METHODS

Study area:

Major cities of the north central of Nigeria, with high population densities [13] and intense industrial activities constitute the area of study. These cities are: Kano, Katsina and Dutse in Kano State, Katsina State and Jigawa state respectively.

Data collection:

Daily data on the waste load weights collected by the trucks of the municipalities were used. To this were added the door-to door and dump -to- dump data collected, because not all parts have access to collection by trucks. Other information was provided by the various States Environmental Protection agencies of the study area. The entire population of the study area selected cities was considered in order to obtain a per capita estimate, according to the National Bureau of Statistics data of 2006. For characterization purposes, the samples were collected during the months of February, March and April and during the raining season in the month of August. The refuse characteristics were then evaluated by sifting through the waste and separate it into its various physical major components.

It should be noted that all samples taken at the various refuse dumping sites contain a large proportion of sand which had to be removed prior to any measurement. This is due to the fact that sand is an inert element in the combustion process. Furthermore, a first sieving is carried out followed by hand picking, carried out to remove small stones which could not pass through the sieve. The sample were then sieved again using a 2 mm mesh sieve, before being ground to powder form using a small porcelain mortar and pestle. Generation rates of MSW were obtained using Equation (1) [3].

$$PCG = \left[\frac{\left(\frac{\text{Waste generated}}{\text{Weeks}} \right) \left(\frac{\text{Weeks}}{\text{Days}} \right)}{\text{Population}} \right] (\text{Kg/ day. person}) \quad (1)$$

where PCG is the per capita waste generation

Materials:

A digital METTLER TOLEDO AB 54 weighting machine with limitations of 10 mg minimum and a maximum of 51 grams. Glass ware equipment include 16 pieces of PYREX conical flasks, pipettes, burettes, beaker, volumetric flask, plastic bottles and filter papers; a digital Spectrometer SPECTRUMLAB 22 PC as well as a Gerhardt - Kjeldatherm machine.

The reagents are: Sulphuric acid (H₂SO₄), Orthophosphoric acid (H₃PO₄), Ferrous sulphate (FeSO₄), Sodium Fluoride (NaF), Potassium dichromate (K₂Cr₂O₇), Diphenylamine indicator, Kjeldhal tabs, Boric acid, Absolute ethanol, Bromolysol green methyl red, Sodium hydroxide. Nitric acid, acetic acid, magnesium sulphate, gum Arabic, barium chloride. Others are Potassium chloride (KCl), Sodium hydroxide (NaOH), Hydrochloric acid (HCl) and Phenolphthalein indicator.

Analytical techniques

Proximate analysis:

This is performed by weighting, heating and burning a small sample of waste to determine the moisture content (M'_{w/o}, in weight percent (w/o)) by driving off the free moisture at ~ 107° C for approximately 1 hour [14] as shown by equation (2).

$$C'_{f(w/o)} = 100 - \left[M'_{(w/o)} + V'_{(w/o)} + A'_{s(w/o)} \right] \quad (2)$$

Ultimate analysis:

It is a quantitative evaluation of the total carbon (C'), hydrogen (H'), nitrogen (N'), sulphur (S'), oxygen (O') percentages after removal of the moisture and ash [15,16]. This analysis is performed using classic oxidation, decomposition, and/or reduction technique to determine, C (carbon), H (hydrogen), N (nitrogen) and S (sulphur). Oxygen O' (w/o) is calculated by difference using the equation as shown below [14]:

$$O'_{(w/o)} = 100 - \left[C'_{(w/o)} + H'_{(w/o)} + N'_{(w/o)} + S'_{(w/o)} + M'_{o(w/o)} + A'_{s(w/o)} \right] \quad (3)$$

Heating value or calorific value:

The ash free, dry heating value can be calculated to 2% accuracy by using the Dulong- Berthelot formula [17]:

$$Q'_d = 81.37C' + 345 \left[H' - \frac{(O' + N' - 1)}{8} \right] + 22.2S \quad (4)$$

All data interpretation, calculations and graphs were carried out and generated using Microsoft Excel, 2007.

III. RESULTS AND DISCUSSION

MSW physical characteristics, generation rates as well as the proximate and ultimate analyses results in each of the towns with higher population density are shown in Figure 1, Table 1 and Tables (2 and 3) respectively.

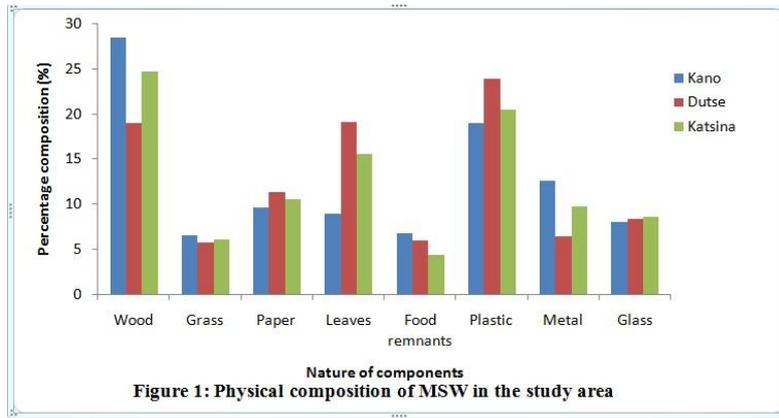


Table 1: Average MSW generation in north central Nigeria

Locality	Generation (kg/person/day)		
	Low	Medium	High
Kano	1.43	1.6	1.8
Dutse	0.68	1.0	1.22
Katsina	1.0	1.12	1.25

Table 2: Proximate analysis of waste samples in north central Nigeria.

Locality	Elements (%)			
	Moisture (w/o)	Volatile matter (w/o)	Ash (w/o)	Fixed Carbon (w/o)
Kano	32.593	25.256	32.150	10.000
Dutse	29.655	20.460	33.260	16.622
Katsina	26.743	23.406	34.750	15.100

Table 3: Ultimate analysis of waste samples in north central Nigeria

Locality	Elements (%)				
	C (w/o)*	H (w/o)	N (w/o)	S (w/o)	O (w/o)
Kano	21.470	0.460	0.685	0.560	12.080
Dutse	21.000	0.120	1.382	0.088	14.495
Katsina	20.965	0.340	1.050	0.291	15.861

* weight percent

The calorific value of refuse (fuel) is the property of fundamental importance. It is a complex function of the elemental composition of the refuse or waste.

Table 4: Calorific value (CV) of waste in the study area

Locality	Calorific value (MJ/kg)
Kano	5.667
Katsina	5.345
Dutse	5.379

The generation and composition of household waste are not homogeneous. They vary according to changes in commercial activities, population behavior, consumption patterns and economic growth rates and depend upon the season of the year, day of the week. Generation rates were found to vary between 0.68 kg and 1.8 kg per person per day. The study are average MSW generation was found to be 1.23 kg per person per day against 1.2 kg per person per day in the north eastern Nigeria [18].

The variation in the CVs of refuse (Table (4)) was due to differences in the gradation of the constituent materials. The refuse CV in Kano with a higher population density approaches the one in Katsina which has a relatively lower

population density. The differences in CVs are not significant even though their population densities are not the same. However, these wastes can immensely impact water quality and sanitation if left untreated.

The present finding is in agreement with Ogaji et al. [12] who found that waste collected from different receptacles and dumpsites in the city of Port Harcourt consisted of 66.6% volatile solids, 13.5% fixed solids, 19.1% liquid and 0.8% other components. They also found an average biodegradability fraction is 0.807, with a carbon-to-nitrogen ratio of 27:1. The energy content of the refuse was 7.25 MJ/kg as collected. Their results indicated that such refuse is amenable to several disposal options with less adverse impact on the environment.

It is well-known that MSW can be used to generate electricity. Research published since the 1970's has reported the use of the biodegradable component in MSW to generate biogas, which can also be used to generate electricity and has positive environmental implications, such as the reduction of greenhouse gas emissions from sanitary landfills and the replacement of highly polluting energy sources (oil, coal and natural gas) [3]. In all the cases, energy recovery is possible with an addition of supplementary fuel since the values in the study area fall below the 7.50 MJ/kg to 12.00 MJ/kg (an acceptable recommended range suggested by Whiting [19]).

Population density and geographic locations are not real determining factors as whether refuse quality may change or not but rather the life style of the population and their level of awareness towards waste management techniques. However, these factors do influence the generation rate of MSW. Options available for supplementary fuels include: Water hyacinth, sugar cane in the Kura, Bunkure, Hadejia areas.

IV. CONCLUSION

Physical characterization showed that wood, grass, metal, plastic and paper were the constituents of all waste samples in the study area, but in varying proportions. Proximate and ultimate analyses of refuse in the area of study showed refuse characteristics as: Moisture: Volatile matter: Fixed carbon: Ash content, as 32.593: 25.256: 10.00: 32.150 for Kano. Average daily municipal solid waste generation was found to be about 1.23 kg per person per day, and it is greatly influenced by population density and commercial activities.

Population density and geographic locations are not real determining factors as whether refuse quality may change or not but rather the life style of the population and the level of awareness of the population towards waste management.

The CVs were low but energy production is possible with the addition of a supplementary fuel such as bagasse, weeds or water hyacinth.

REFERENCES

- [1] Daskalopoulos E., Badr O. and Probert S., An integrated approach to municipal solid waste management, *Journal of Resources, Conservation and Recycling*, 24(1), 1998, 33-50.
- [2] Ehrlich Paul R, and Ehrlich Anne H., *Population, Resource, and Environment: Success in Human Ecology* (H. Fressman and Co. San Francisco 2nd edition, 1996).
- [3] Aguilar-Virgen Q., Armijo-de Vega C., P.A. Taboada-González and S. Ojeda-Benítez, *Municipal Solid Waste Generation and Characterization in Ensenada, Mexico; The Open Waste Management Journal*, 3, 2010, 140-145.
- [4] Syngenta Foundation for Sustainable Agriculture (SfSA), *Sustainable innovation in agriculture: A Background to the Sahel and the semi-arid region* (Helika Limited, UK, 2007).
- [5] Ikuponisi Francis S., *Status of Renewable Energy in Nigeria, An international Conference on making renewable Energy a Reality. ONE SKY- Canadian Institute of Sustainable Living*, 2006, 21-27.
- [6] Diso I. S., *Waste Management Survey in Kano, The Energy Team Bayero University Kano, Report submitted to the Kano State Government*, 1995.
- [7] Mengiseny E. Kaseva and Josia L. Moirana, *Problems of solid waste management on Mount Kilimanjaro: A challenge to tourism. Waste management and Research*, 28(8), 2010, 695-704.
- [8] Agunwamba J. C., Ukpai O. K. and Onyebuanyi I. C. *Solid waste management in Onitsha, Waste management and Research*, 16 (1), 1998, 23-31
- [9] Adeyemi A. S., Olorunfemi J. F. and Adewoye T. O. *Waste scavenging in third World cities: A case study of Ilorin, Nigeria, The Environmentalist*, 21(2), 2004, 93-96. Springer Netherlands.
- [10] Jekayinfa, S. O. and Omisakin, O. S. *The energy potentials off Some Agricultural Wastes as local Fuel materials in Nigeria, Agricultural Engineering International: the CIGR Ejournal. Vol.VII. May 2005, Manuscript EE 05 003.*
- [11] Paolo S. Calabro, *The effect of separate collection of municipal solid waste on the lower calorific value of the residual waste. Waste management and Research*, 28 (8), August 2010, 754-758.
- [12] Ogaji S. O. T., S.D. Probert, A.H. Igoni and M.J. Ayotamuno, *Municipal solid-waste in Port Harcourt, Applied Energy*, 84(6), 2006.
- [13] National Bureau for Statistics, *The National Population Census. Annual report, Nigeria, 2006, 01-07*
- [14] Weisman Joel and Roy Eckart, *Modern Power Plant Engineering* (2nd edition, Prentice- Hall International, New Delhi, 1986).
- [15] Tabatabai M. A., *Determination of Sulphates in Water Samples, Sulphur Institute Journal*, 10, 1974, 11-13.
- [16] Page A. L., R. H. Miller and D. R. Keeney, *Methods of Soil Analysis, Handbook of Chemical and Microbiological Properties –Part 2 (Second Edition, 1982).*
- [17] Lee C. C. and Huffman G. L., *Solid Waste Calculations: Thermodynamics used in Environmental Engineering, Handbook of Incineration; 2 (U.S. Environmental Protection Agency, Cincinnati, Ohio, 2001), 2.47.*
- [18] Oumarou M.B., Dauda M., Sulaiman A.T. and Babagana M.T., *Characterization and Generation of Municipal Solid Waste in North eastern Nigeria; Continental Journal of Renewable Energy*, 2012, Ref: Cjre/2012/020
- [19] Whiting K., *Large scale MSW incineration Technologies. In: Annual short course on Incineration of municipal Waste Report, University of Leeds, Department of Fuel and Energy, Leeds, UK, 2002), 1-37.*

An Enhanced Protocol for Solving Broadcast Storm Problem in MANET

K.B.Gurumoorthy¹, K. Manojkumar²

Asst.Professors, Department of ECE, Christ the King Engineering College, Coimbatore,

Abstract: Mobile Ad-Hoc Network (MANET) is wireless networks consisting of a collection of mobile nodes with no fixed infrastructure, where some intermediate nodes should participate in forwarding data packets. Energy conservation is a critical issue in ad hoc wireless networks for node and network life. Enhanced Medium Access control (eMAC) protocol prevents link/routing failures, hidden/exposed terminal problems and broadcast storm problems using an adaptive unreachability reporting mechanism with more energy consumption. Furthermore, An adaptive table broadcasting technique is proposed to distribute topology information in mobile ad hoc networks (MANETs). In this paper, a cross layer design for enhancing the distance based broadcasting protocol is proposed in terms of energy consumption. Instead of using the distance, the reception signal strength is considered. The necessary transmission power to reach an intended device is obtained using the beacons. If the furthest node can be reached using less power than the default value, the transmission power is reduced and it saves energy. Different proposals for enhancing the algorithm are proposed, and they not only save energy but also highly reduce the number of collisions.

Keywords: Energy efficiency, Mobile Ad-hoc Networks, cross layer design, unreachability and distance based broadcasting.

I. INTRODUCTION

Mobile ad-hoc network (MANET) is composed of clusters of self-organized wireless stations without a need to utilize any preinstalled infrastructure. Due to the prospective of self-organized deployment, lots of practical applications have been conceived for MANETs and the efficiency of MANETs depends on the performance and reliability of the medium access control (MAC) protocol applied in such environments.

The unreachability problem becomes more severe in multihop environments and results in packet dropping, starvation of part of traffic flows, and possibly unnecessarily network-layer rerouting [2]. The protocol in [1] adds a couple of new control frames to ease the reporting of the unreachability situation to solve the receiver-blocking problem. When a station is notified about an upcoming data communication due to which it will be unreachable, it is given an opportunity to inform its one-hop neighbors about the forthcoming unreachability. In principle, right after the RTS/CTS negotiation and before commencing the actual DATA transmission phase, the stations, which will shortly become unreachable, are given the chance to report their imminent unreachability status using a designated broadcast frame called individual communication pause (ICP). Collisions may occur among broadcasted ICP frames. Such collisions are caused by unconditional ICP frame broadcasting, it refers to it as ICP broadcast storm. Here, the broadcast storm problem is solved by introducing a

technique to prevent unnecessarily simultaneous unreachability reports and maintenance of a double-hop neighborhood (DHN) graph by every station. The DHN graph of each station gives an estimate on its DHN topology. By incorporating topology-awareness and smarter decision-making algorithms into the MAC protocol, the impact of the unreachability problem is reduced, resulting in much more efficient channel utilization and higher transmission capacity.

One of the main problems in dissemination is the broadcast storm problem [7]. Not only the problem was presented in [7] but some different techniques for minimizing its effects were introduced, like (1) the probabilistic scheme where nodes resend the message with a predefined probability. (2) The counter based approach that forwards the message in terms of the number of copies received. (3) The distance based technique that considers candidate nodes for forwarding those further from the source than a predefined distance. (4) In the location based approach, the receiver knows the position of the source, so it is able to calculate precisely the additional area covered with the forwarding. (5) Finally, in the cluster based scheme, nodes are distributed in clusters. Only those nodes considered as head or gateway are candidates for forwarding. All these approaches try to minimize the number of forwarding nodes. In this scheme, the distance based broadcasting protocol is considered, that aims at selecting forwarding nodes in terms of the distance between the receiver and the source node, and enhancing it by minimizing the transmission power every node uses for the broadcasting process in order to save energy and reduce the number of collisions.

Energy consumption is more important aspect because ad-hoc networks are composed of devices that rely on batteries. Reducing the transmission power will directly increase the battery life of the nodes and thus, the network lifetime. The contributions of this paper are two folds: (1) adding energy efficiency features to the distance based approach by reducing the transmission power of the source nodes, (2) analyzing the influence that reducing the transmission power has over other nodes in terms of the number of collisions or the interference level.

II. RELATED WORK

A. Introduction

Literature survey is carried out by analyzing many papers relevant to unreachability problem like hidden/exposed terminal problems and distance based approach to reduce energy consumption of nodes in MANETs. The researches carried out by different authors are surveyed and the analysis done by the researchers are discussed in the following paragraphs.

B. Dual Busy Tone Multiple Access -A Multiple Access Control Scheme for Ad-Hoc Networks

Z. J. Haas and J. Deng [3], proposed that the dual-BT multiple access (DBTMA), where two out-of-band Busy

tones (BTs), i.e., BT_t and BT_r are deployed to protect RTS and DATA frames, respectively, the former is always activated by the source station when an RTS frame is being transmitted, while the latter is triggered by the destination station when it is receiving a DATA frame. In contrast, DUCHA utilizes two channels separately for control and data frames. RTS and CTS are transmitted in a separate control channel to avoid collisions with the data packets. Negative Clear to send (CTS) is used to solve the receiver-blocking problem and is also transmitted in the control channel. An out-of-band receiver-based BT is used to solve the hidden-terminal problem. In ad hoc networks, the hidden and the exposed terminal problems can severely reduce the network capacity on the MAC layer. To address these problems, the ready-to-send and clear-to-send (RTS/CTS) dialogue has been proposed in this literature. However, MAC schemes using only the RTS/CTS dialogue cannot completely solve the hidden and the exposed terminal problems. A new MAC protocol DBTMA scheme operation is based on the RTS packet, two narrow bandwidth and out of band busy tones.

C. AMACA—A New Multiple Access Collision Avoidance Scheme for Wireless Lans

K. Ghaboosi and B. H. Khalaj [1], proposed that “hidden-and exposed terminal” are among the main problems in ad-hoc WLAN networks. In addition, there are scenarios where the desired destination is located in the range of other transmitters, so that the efforts on setting up communication with this terminal will fail due to collisions occurred between desired control packets and unwanted received data packets at destination. In such scenarios, conventional protocols can not address the problem efficiently, resulting in throughput and channel utilization degradation. By using the same PHY of IEEE 802.11 and making slight changes in its MAC layer, a new MAC protocol is presented to address such problems. The performance of this method is better than IEEE 802.11 and DBTMA besides, in addition to solving the above problems, it improves channel utilization and reduces the total overhead. AMACA is a MAC-only modification based on the IEEE 802.11 standard. In this scheme, the issue of QoS is addressed by *Prior Channel Reservation* feature. Due to the modified features, the total overhead is reduced, resulting in higher channel utilization. In addition, the resulting average jitter delay of received data in AMACA is expected to be less than the other protocols.

D. A Distributed Energy-Efficient Routing Algorithm Based on cross layer design

F. Zheng, H. Lu, W. Wang and Q. Sun [12], proposed a distributed energy-efficient routing algorithm for mobile Ad Hoc networks (MANETs). The cross-layer design paradigm is adopted. The distance from the source node to the destination node is estimated based on the received signal strength indication (RSSI) of the packets and is used to adaptively adjust the backoff time of the MAC layer. The distance threshold and a packet count threshold are used to schedule the transmissions of packets in the Network layer. The algorithm is distributed and works without needing any global network information or control packet. The results of this method show that the routing algorithm is energy-efficient and drastically alleviates the Broadcast storm problem. For coping with the Broadcast Storm Problem, a

distributed energy-aware routing algorithm is presented for MANETs. The algorithm selects rebroadcast nodes based on the received signal strength information extracted from the physical layer and the nodeneighboring information extracted from the network layer. It adaptively adjusts the backoff time in the MAC layer and the packet delay time in the network layer.

III. ENHANCED MEDIUM ACCESS CONTROL (eMAC)

For reducing the unreachability problem, the general architecture of eMAC protocol is chosen and it is analysed. The main aim is to introduce a technique to avoid the spread of unnecessarily simultaneous unreachability reports, and therefore, the Double Hop Neighbourhood (DHN) graph is maintained by every station. Each station can be either mobile or stationary. The DHN graph of each station gives an estimate of its DHN topology. This may be accomplished by overhearing Request to Send / Clear to Send (RTS/CTS) control frames if the unreachable station is situated in the communication range of the unreachability cause (i.e., unreachability of type I). On the other hand, for those stations situated in the interference range of the unreachability cause but not its communication range (i.e., unreachability of type II), this goal may be achieved by overhearing individual communication pause (ICP) frames of type I received right after a BT of a particular duration.

In this case, the unreachable station of type II broadcasts an ICP frame of type II for which there are two address fields: The first address field carries the MAC address of an unreachable station of type I from which an ICP frame of type I has been received, and the second address field carries the MAC address of the unreachability cause. In this architecture, an ICP frame generated by an unreachable station of type I is referred to as an ICP frame of type I. Similarly, an ICP frame generated by an unreachable station of type II is referred to as an ICP frame of type II. The former has only one address field used for carrying the MAC address of the unreachability cause, while the latter has two address fields. For simplicity, an ICP frame of type I is denoted by ICPv1. Similarly, an ICP frame of type II is denoted by ICPv2.

In addition, its duration/ID field is used to indicate the duration of unreachability. The announcement of an upcoming unreachability status is performed either right after an overheard RTS and/or CTS frame (unreachability of type I) or upon overhearing an ICP frame of type I received right after a BT of a particular duration (unreachability of type II). Basically, the eMAC table is generated from the DHN graph. On the other hand, upon reception of all one hop neighbors' eMAC tables, each station either constructs or updates its local DHN graph as well. To clarify this issue, consider the network topology illustrated in Fig.1. In this configuration, station “A” can receive eMAC tables from all its immediate one hop neighbors, i.e., stations “B,” “C,” and “F.”

Similarly, station “B” is able to obtain tables from stations “A,” “F,” “G,” and “H.” Station “A” is able to easily construct a DHN graph to mimic its DHN topology. Now, Assume that stations “G” and “H” are willing to perform a long-term data exchange using packet fragmentation. In this scenario, “G” is supposed to serve as the source station, and “H” is assumed to be the destination station.

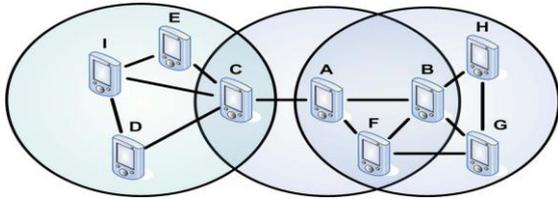


Fig.1. Exchange of Topology Information using eMAC.

Apparently, this leads to the unreachability of both “B” and “F.” When station “A” receives an ICP frame sent by either “B” or “F,” it verifies the appended field indicating the unreachability cause and using its DHN graph, it concludes that all one-hop neighbors of station “G” will be unreachable as well. Here, “A” needs to receive only one ICP frame to be informed about the unreachability of “B” and “F.” For the case of an unreachable station of type II, the same approach is followed by Station “A.” As in ICP frames of type II, both the MAC address of an unreachable station of type I, from which an ICP frame of type I has been received, and the MAC address of the unreachability cause are appended station “A” concludes that not only “B” and “F” but that any one-hop neighbor of these two stations will be unreachable as well. The Fig.1. shows exchange of topology information using eMAC table.

A. DHN Graph

The DHN graph is a time-variant data structure denoted by $G(V(t), E(t))$, where $V(t)$ represents the set of its vertices, and $E(t)$ stands for the set of its edges at a given time instant t . Based on topology information received from one-hop neighbors in the form of eMAC tables, the DHN is accordingly updated. When the DHN graph is updated, the synchronous eMAC table, which is denoted by $\Xi(t)$, is generated. Each station records all its neighbors in its local DHN represented by a set of vertices classified into two different categories. One-hop neighbors are grouped together to form the class-N group. Each member of this group is simply referred to as a class-N neighbor. Fig.2. illustrates the DHN graph maintained by station “A”

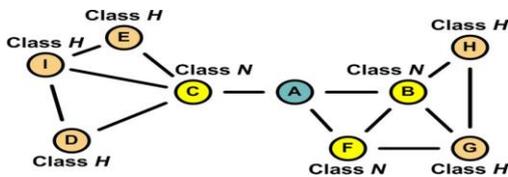


Fig.2. DHN Graph Maintained by station ‘A’.

In addition, double-hop neighbors are grouped into another set named as a class-H group, each member of this group is class-H neighbour. Principally, in the DHN graph, there should be a unique edge connecting each class-N neighbor to the local station, which is the owner of the DHN graph. On the other hand, class-H neighbors are connected to the local station via a couple of edges and obviously through class-N neighbors.

B. eMAC Table Structure

Each station maintains two different versions of the eMAC table at any time. One is called the synchronous eMAC table and is denoted by $\Xi(t)$. This table is directly generated from the local DHN, whenever it is updated. The second table is denoted by Ξ and represents the latest version of the eMAC

table that has been broadcasted over the air interface. Basically, each station should broadcast its synchronous eMAC table $\Xi(t)$ in a regular fashion; whenever the synchronous eMAC table $\Xi(t)$ is broadcasted in its corresponding due beacon interval (BI), the existing Ξ table is simply replaced by $\Xi(t)$.

C. eMAC Table Maintenance and Broadcasting Rules

When the DHN graph is updated due to the reception of new neighborhood topology information, the synchronous eMAC table $\Xi(t)$ is consequently regenerated. This means that the DHN and $\Xi(t)$ keep the most up-to-date information about the DHN and one-hop neighborhood of the local station, respectively. As stated earlier, the most up-to-date version of the eMAC table should be broadcasted in a regular fashion. To determine how frequent and when the eMAC tables are broadcasted, the number of BIs that have to elapse before broadcasting the latest version of the eMAC table is specified. When the synchronous eMAC table $\Xi(t)$ is broadcasted in its due BI, it is also saved as Ξ to represent the last version of the local eMAC table that has been broadcasted over the air interface.

IV. DISTANCE BASED (DB) BROADCASTING ALGORITHM

Distance Based (DB) is one of the different schemes proposed for minimizing the effects of the broadcast storm problem when disseminating information in wireless networks. The protocol makes use of the distance between the source node and the receiver. The idea is that a node receiving a broadcast message for the first time will compute the distance to the source node. If this distance is small, the contribution to the dissemination performing this forwarding is negligible and therefore, the message is not resent. Only nodes that are separated at least a minimum distance from the source node resend the message. This minimum distance is a predefined threshold, D . The protocol also includes a delay before forwarding a received message, and if the same message is heard more than once (during this waiting time), the delay is cancelled.

Fig.3. represents the functioning of the algorithm. Considering node A broadcasts a message m , nodes B and C will not resend m because the distance from those nodes to A is smaller than D . Nodes E, F and G will wait for a random number of slots. If node F finishes the waiting time first, it will forward the message and, thus, node E will hear it and calculate the distance from node F. as the distance is smaller than D , node E will the packet

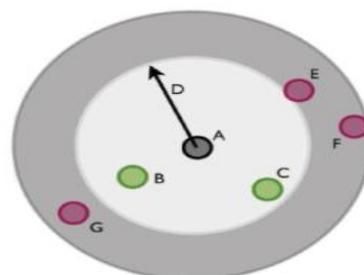


Fig.3. Mechanism of DB.

A. Enhanced DB

In this section we explain the procedure followed to implement the broadcasting algorithm, and also the improvements introduced to the original protocol, DB.

1) Implementation

For calculating the distance between a source and a destination, the signal strength of the received packets is used to estimate how far two nodes are. In this implementation, the threshold D is not in terms of distance (m) but power (dBm). It is called *borders Threshold* as it defines the nodes that are considered to be far from the source and therefore close to the border. The value used for this parameter is -90 dBm. This value was experimentally chosen, and any value below it supposes that the source and destination nodes are separated at least $2/3$ of the maximum coverage. A node is not able to decode a received packet if the reception power is lower than -95 dBm, this is called the *end Threshold*. Therefore, all nodes whose reception energy vary from $[-95, -90]$ dBm are candidates of forwarding the broadcasting message. Every device sends a *hello* message (or beacon) to alert devices within a close area about their presence. A device receiving these beacons is able to keep track of all neighbors around. In this situation the algorithm is able to take decisions depending on this value. When a broadcast message is sent, the receiving node will check the reception power, if it is below the *borders Threshold* (-90 dBm), it will consider itself as a bordering node and thus, sets the delay.

2) Enhancements

An ad hoc networks, and devices depend on battery, saving energy supposes one critical aspect. One of the new features added to DB is reducing this energy consumption using transmission power reduction.

Reducing Transmission Power

In any wireless transmission, as the electromagnetic wave propagates through the space, the power of the signal suffers from path loss attenuation causing a reduction in the signal power. The relation between the transmitted power and the power finally received at the destination in terms of dB is expressed as,

$$\text{Received Power} = \text{transmitted Power} - \text{loss} \text{ ----- (1)}$$

Thus, a node receiving a beacon will be able to estimate the loss that packet suffered during the transmission, using the reception power detected at the physical layer. Every node keeps and updates the reception power of each of its neighbors in a list. When a device wants to send a broadcast message, it will be able to estimate the the packet loss. If a node can estimate the loss the packet is going to suffer, it will be able to reduce its transmission power and use only the necessary one to get the furthest one hop neighbor. Thus, reducing the transmission power for sending broadcast messages directly decreases the energy consumption of the device, without degrading the performance of the broadcasting process. so that the furthest node is receiving the packet with the minimum reception power allowed to correctly decode the message. That means, its reception power should be the *end Threshold*. the new reduced transmission power can be calculated as,

$$\text{Transmission Power} = \text{loss} + \text{end_Threshold} \text{ -----(2)}$$

From the above equation, it is possible to estimate the maximum transmission power needed to reach the furthest neighbor in the one hop neighborhood. If it is less than the default transmission power, It is reduced in order to save energy, Therefore, reducing the transmission range from r to r' decreases the energy consumption with no detriment of the network connectivity as shown in Fig.4..

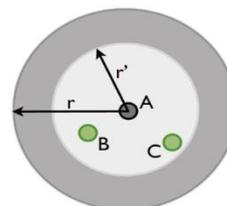


Fig.4. Reducing the Transmission Power of Nodes.

Reducing the transmission power for sending broadcast messages not only improves the energy consumption in wireless networks, but also reduces the interference level of devices in a close area. Each device has the *end Threshold* from which on, if the received signal strength is lower, the device will not be able to recover the data transmitted, but this reception will be considered as noise and will increase the interference level of the device.

Using Different Delay Techniques

DB stops the random delay when a repeated message is heard. Then, if the distance from the new source node is smaller than the threshold D , the message is discarded and no retransmission is performed. Otherwise, the forwarding starts. Instead of stopping the delay when a repeated message is heard, the possibility of keeping tracks of the received energy and continue the delay are considered.

In this section, the behavior of two different techniques are considered and comparing them to the original proposal of DB,

1. In the first one a fixed delay inversely proportional to the received power is considered. the procedure to calculate the delay in terms of the reception power is shown as,

$$\text{Power Delay} = -1/\text{rxPower} - \text{borders_Threshold} - 1 \text{ ---- (3)}$$

2. The second proposal considers a random delay chosen from an interval whose size also varies with the reception power. That is, the waiting time will be chosen between $[0, \text{power Delay}]$ and the delay varies from 0 to 1 second.

V. SIMULATION RESULTS

To evaluate the performance of proposed protocol, extensive simulations are provided for the system throughput, delay, jitter, and overhead and compare the achieved results with Enhanced MAC (eMAC) protocol. The propagation model is the two-ray ground model, the transmission range of each station is approximately 250 m, the carrier sensing range is approximately 400 m. The channel rate is set to 2 Mb/s and mobile nodes exist in an area $2,500$ [m] x $2,500$ [m]. In this simulation study, the following performance metrics are evaluated.

A. End-to-End Delay

The term the average delay is a data packet experiences to cross from source to destination. This delay includes all possible delays caused by buffering during route discovery delay, queuing at the interface queues, propagation and transfer times.

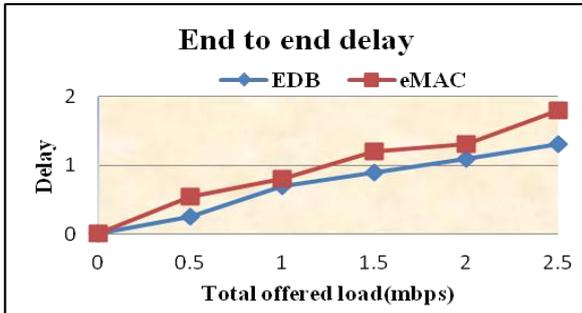


Fig.5. Total offered load Vs Delay.

The Fig.5 shows the comparative delay analysis for eMAC and Enhanced Distance Based (EDB) protocol on varying the offered load. The graph shows that eMAC protocol has less end to end delay compared to EDB protocol even if the end to end delay increases.

B. Normalized Overhead

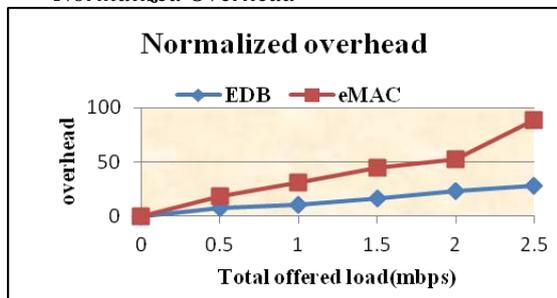


Fig.6. Total offered load Vs Overhead.

This overhead is the ratio of number of routing control packets to delivered data packets. Each packet requires extra bytes of format information are stored in packet header, which reduces the overall transmission speed of raw data. The above Fig.6 illustrates the normalized control overhead for two different MAC protocols versus the total offered load. The proposed EDB shows indeed better results in comparison with eMAC scheme.

C. Jitter

This jitter is defined as the variation in the packet delay. High jitter means the difference between delays is large and low jitter means the variation is small.

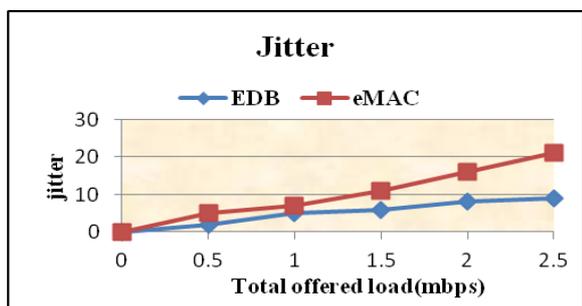


Fig.7. Total offered load Vs Jitter.

Fig.7 shows the average frame jitter versus the total offered load in megabits per second. The EDB protocol shows the best performance compared with eMAC protocol when the offered load varies between 0.5 and 2.5 Mb/s.

D. Energy consumption

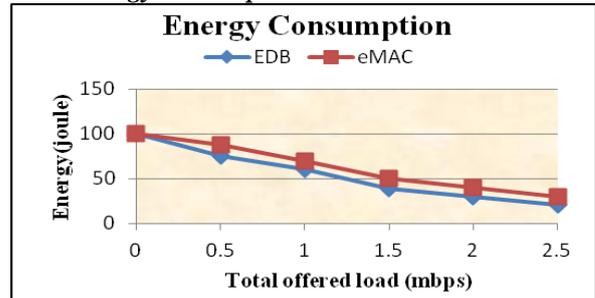


Fig.8. Performance of energy consumption across various loads.

Fig.8 presents the energy consumption. The Comparison of energy consumption for EDB with that of the eMAC protocol is shown. It is clearly seen that energy consumed by Enhanced DB protocol is less compared to other schemes.

E. Throughput analysis

The term throughput is the ratio of the total amount of data that a receiver receives from a sender to a time it takes for receiver to get the last packet. A low delay in the network translates into higher throughput.

One-hop throughput

One hop throughput is the number of data packets transmitted between two successive nodes.

Fig.9 illustrates the one hop throughput for different schemes when the total offered load varies. It shows that EDB has high throughput compared to eMAC protocol even if the one hop throughput increases.

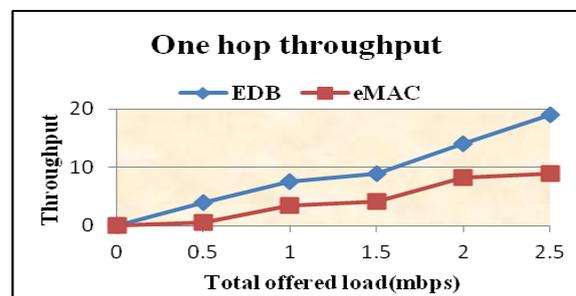


Fig.9. One-hop throughput Analysis.

End-to-End throughput

The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.

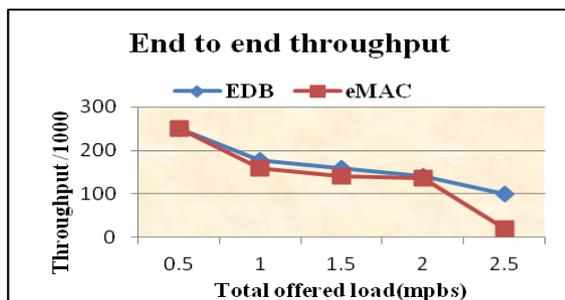


Fig.10. End-to-end throughput Analysis.

Fig.10 shows the aggregate end-to-end throughput versus the total offered load in megabits per second. eMAC shows the worst performance compared with EDB protocol when the offered load varies between 0.5 and 2.5 Mb/s.

VI. CONCLUSION

The unreachability problems have addressed here without deployment of more than one communication channel. The proposed eMAC protocol scheme prevent the unreachability problem, resulting in much more efficient channel utilization and higher transmission capacity by implementing topology-awareness and smarter decision-making algorithms into the MAC protocol. It has been shown that the unreachability problem can be addressed in a better way, leading to an adaptive and robust topology-aware protocol with more energy consumption. An energy saving strategy for the well known distance based broadcasting algorithm DB is proposed.

For decreasing the energy consumption, a reduction in the transmission power is performed when possible. This is really useful when the network is not very dense reducing up to 86.97% in the best case, but when the number of devices is big, the node does not reduce the transmission power so much since there are usually nodes close to the border. This strategy of reducing the transmission power is saving at least 7.55% of energy per forwarded message. As a result from the experiments performed, The enhanced distance based (EDB) protocol is the one that generally behaves better than the enhanced medium access (eMAC) protocol. In this work enhanced distance based protocol is able to reduce energy without degrading the network connectivity and that also reduces the number of collisions in a 95.41%. The simulation results have showed that the EDB protocol has better performance than enhanced medium access protocol in terms of end to end delay, jitter, throughput analysis and overhead.

REFERENCES

- [1] K. Ghaboosi and B. H. Khalaj, "AMACA—A new multiple access collision avoidance scheme for wireless LANs," in *Proc. IEEE PIMRC*, 2004, pp. 238-242.
- [2] H. Zhai, J. Wang, and Y. Fang, "DUCHA: A new dual-channel MAC protocol for multihop ad hoc network," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3224-3233, Nov. 2006.
- [3] Z. J. Haas and J. Deng, "Dual busy tone multiple access (DBTMA) a multiple access control scheme for ad-hoc networks," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 975-985, Jun. 2002.
- [4] Y. Li, H. Wu, D. Perkins, N. Tzeng, and M. Bayoumi, "MAC-SCC: Medium access control with a separate control channel for multihop wireless networks," in *Proc. ICDCSW*, 2003, pp. 764-769.

- [5] V. Shah, S. Krishnamurthy, and N. Poojary, "Improving the MAC layer performance in ad hoc networks of nodes with heterogeneous transmit power capabilities," in *Proc. IEEE ICC*, 2004, pp. 3874-3880.
- [6] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Solutions to hidden terminal problems in wireless networks," in *Proc. ACM SIGCOMM*, 1997, pp. 39-49.
- [7] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in *ACM/IEEE Int. Conf. on Mobile computing and networking*, 1999, pp. 151-162.
- [8] C. Zhi-yan, J. Zhen-zhou, and H. Ming-zeng, "An energy aware broadcast scheme for directed diffusion in wireless sensor network," *Journal of Communication and Computer*, vol. 4, no. 5, 2007, pp. 28-35.
- [9] X. Chen, M. Faloutsos, and S. V. Krishnamurthy, "Power adaptive broadcasting with local information in ad hoc networks," in *Conference on Network Protocols*. IEEE Computer Society, 2003, p. 168.
- [10] D. Liarokapis, A. Shahrabi, and A. Komninos, "Diba: An adaptive broadcasting scheme in mobile ad hoc networks," in *Communication Networks and Services Research*, 2009, pp. 224-231.
- [11] J. Gomez and A. T. Campbell, "Variable-range transmission power control in wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 1, 2007, pp. 87-99.
- [12] F. Zheng, H. Lu, W. Wang and Q. Sun, "A Distributed Energy-Efficient Routing Algorithm Based on cross layer design", in *Conference on National Natural Science*, 2009, pp. 197-200.
- [13] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE 802.11 WG, Aug. 1999.
- [14] K. Ghaboosi, M.L Aho, Y. Xiao and Q. Zhang "eMAC A Medium Access Control Protocol for the Next-Generation Ad Hoc Networks", *IEEE Transaction on vehicular technology*, vol.58, no.8, pp 4476-4490, 2009.
- [15] A. Boukerche, *Handbook on Algorithms for Wireless Networking and Mobile Computing*. London, U.K.: Chapman & Hall, 2005.

A Classification of Job Scheduling Algorithms for Balancing Load on Web Servers

Sairam Vakkalanka

School of computing, Blekinge Institute of Technology, Karlskrona, Sweden-37141

ABSTRACT: Through this report, a classification of different job scheduling algorithms available for balancing the load on web servers is made. Types such as static and dynamic scheduling algorithms are thoroughly discussed and the strengths and weaknesses of these algorithms are put forth through this article.

Keywords: Load balancing, scheduling algorithms, web servers, traffic, load index

I. INTRODUCTION

With the rapid increase and growth of World Wide Web (WWW), grew the usage of several complicated and computation-intensive applications, which require high degree of computation and higher bandwidth for the transmission of data [26]. These applications may vary from cloud based, multimedia, design and development, e-commerce etc [25]. With these options being made available for users all over the world, there is an exponential increase in the usage of network bandwidth. This increase or change is not only affected by the traffic but also by the nature of traffic, which in the era where web servers were used for the first time were used only to transfer plain texts or images [25]. Now, with the explosion of data, traffic and low bandwidth problems, balancing the load on these web servers play a vital role.

II. TRAFFIC AND ITS TYPES

As stated earlier, load on these web servers not only depends on the traffic but also on the type of traffic. According to Kotogiannis et.al [13], traffic on these web servers can be classified into

- **General traffic**
- **Secure traffic**
- **Multimedia traffic**
- **Burst traffic**
- **Non congestive traffic**

General traffic:

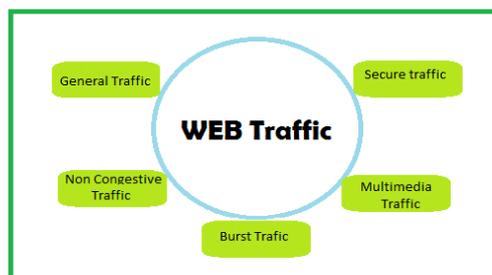
This sort of traffic can be stated as the traffic generated due to request for data such as the plain text documents or static content on web pages and dynamic content [13].

Secure traffic:

This type of traffic is mostly generated by e-commerce applications, which largely run on the SSL- TTL protocol [18].

Multimedia traffic:

The multimedia traffic is a sort traffic which is generated by the streaming of data which may either be video or audio [18]



Non congestive traffic:

Though this sounds like general traffic, it is distinguished in terms of the size of the packet [13][20]. The packet size in a non congestive traffic is usually small (NCQ threshold)[13][20]. This kind of traffic never leads to jitter or delay [13][20].

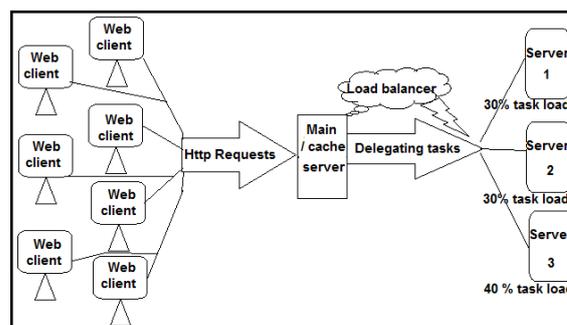
Burst traffic:

This type of traffic is mainly caused due to packet which are transferred in bursts such as P2P transfers and file downloads or uploads etc [16][13].

With these different types of traffic exist different load balancing techniques. These load balancing techniques and their types are discussed in the section below.

III. LOAD BALANCING

Load balancing is used to distribute work between two or more processors, computers, networks or memory devices in order to channelize the resources in an efficient manner and to get optimized response times and throughputs [1]. Load balancing can be defined as an approach to increase and improve the performance of two or more nodes or links connected nodes by the redistribution or the reassignment of load [6][9][10]. The figure below explains how load balancing works in a web.



A. Main Goals of load balancing

According to [6][11], Balancing the load on the nodes and links in a distributed setting is always driven by the goals discussed below

- To provide, a plan B when a single node or group of nodes fail.
- To improve the overall performance of the connected nodes or network.
- To maintain the stability of the systems connected.
- To make available systems for easy future modifications.

This load balancing is always fruitful and has many advantages when the goals are satisfied. The advantages of load balancing are discussed in the following section.

B. Features and advantages of Load balancing

Balancing the load on servers comes with added features and benefits though increases the cost of communication and transfer between the nodes. Some of those advantages and features are listed below:

- Load balancing protects the servers from Distributed denial of Service attacks (DDos)
- Balancing the load improves the reliability of systems, reducing the crashes on the nodes caused due to overload.
- Load balancers can help buffer response from the servers and slowly send to the clients who are down, reducing the burden and waiting time on the servers.
- Load balancers have the feature of asymmetric load distribution where overloaded tasks can be assigned to servers at the backend.
- Load balancing helps in improving functionality, stability, reliability and maintainability of the servers.

Load balancing can be considered as a process which is carried out in such a way that no processes are overloaded but kept busy [1]. In order to know if a node is busy or not and to check the load on the node, Load index is calculated.

C. Load Index

Load index is used to identify or to detect an imbalance state [1]. An imbalance state occurs when the load index of a particular node is greater than the load indices of others which vary with a variation in the performance measure of interest [1]. The performance measure of interest can be anything, for example the Length of the CPU queue can be considered when the performance measure of interest is the average response time [1][3][4]. All load balancing algorithms are based on this load index and also some governing policies which are discussed below.

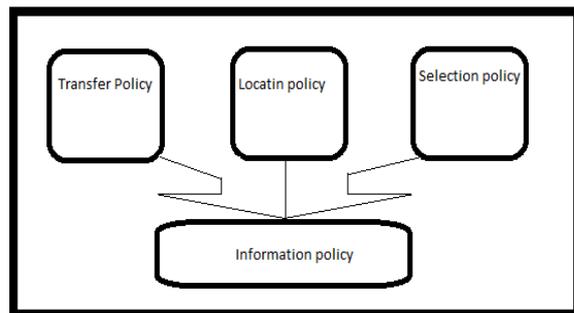
D. Load balancing policies

All load balancing algorithms are based mainly on four policies, which are responsible in keeping the systems updated with the information of workload on the nodes [1]. The four policies which govern the load balancing algorithms are as follows

- Global Information Policy
- Transfer Policy
- Location Policy
- Selection Policy

Information Policy gives all the nodes an access to the load indices of each and every node, which comes with an added cost of extra effort needed for communication in order to maintain the exact information of the nodes[1][2][5][6]

Transfer Policy determines when a node can distribute the load or transfer a job to the other node, also when a node can receive the load or retrieve a job from another node [1][6]. A node becomes eligible to transfer or receive when it crosses or reaches a certain threshold limit which is determined by the total average load on these nodes [1][6]

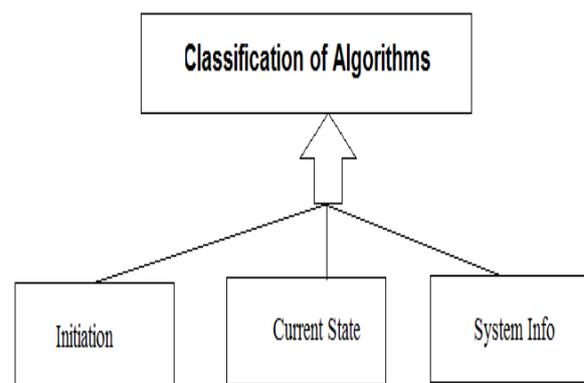


Location Policy determines which node needs to be paired with another in order to accomplish the transfer of load or job [1]. If the node is a sender then location policy looks for a receiver and vice versa [6].

Selection Policy selects the appropriate jobs from the queued jobs in order to retrieve / transfer the task to an eligible receiver / sender [1]. This policy works on the principle of minimizing the cost required to transfer the jobs from one node to the other [1][6]

IV. SCHEDULING ALGORITHMS FOR LOAD BALANCING

The main aim of scheduling algorithms is to improve the stability, reliability and performance of systems which are connected in a network. There exist different kinds of scheduling algorithms which are explained below:



Classification of scheduling algorithms in load balancing can be done in three ways as explained by different authors are as follows

- Classification based on Initiation
- Classification based on system information
- Classification based on state of the current system

E. Classification based on Initiation

Here, scheduling algorithms are classified based on the job transfer initiation process [6][11].

- Sender initiated algorithms
- Receiver initiated algorithms
- Symmetric algorithms
- If sender initiates the process, then the algorithms pertaining to the sender are considered as sender initiated algorithms [6][11].
- If the receiver initiates the process, then the algorithms which fall under this category are considered to be receiver initiated algorithms [6] [11].
- If both sender and receiver simultaneously initiate then they are considered to be symmetric algorithms [6] [11].

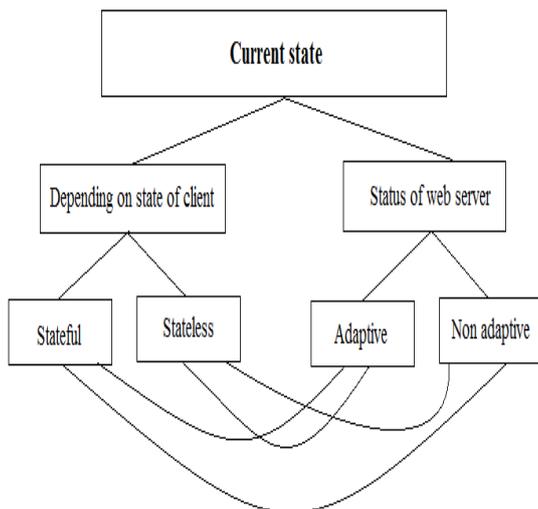
F. Classification based on state of current system

Depending on the state of the systems, load balancing algorithms can be classified into two ways

- ❖ Depending on the state of client request
- ❖ Depending on the status of the web server

▪ **Depending on the state of client request**

If algorithms need information regarding connection requests made by nodes or clients connected in a network [13], then they are classified into



State full Algorithms

These are those algorithms which require the information regarding connection requests made by the nodes [13].

Stateless algorithms

These are those algorithms which do not require the information regarding the connection requests made by the nodes [13].

▪ **Depending on the status of the web server**

Based on the status of the server [13], algorithms can be classified in to two ways

Adaptive algorithms

These are those algorithms which require the status of the server [13].

Non adaptive

These are those algorithms which do not require the status of the server [13].

These are again combined into four categories namely,

- Stateless non adaptive
- State full adaptive
- State full non adaptive
- Stateless adaptive

▪ **Stateless non adaptive**

These algorithms do not take into regardsystem information where it may be the client connection status or the status of the web server [13]. Algorithms such as Random and round robin algorithms come under this category stateless non adaptive algorithms [13][15][19].

▪ **State full adaptive**

These are those algorithms which make use of information from both servers and nodes, which is based on the ratio of Number connection requests at a node to the average connection requests received with a particular time interval [13].

$$t_2 - t_1 : R_i = |ci_2 - ci_1| / \frac{1}{n} \sum_{i=1}^n Ci_1 .. [13][21][24].$$

Least loaded algorithm which falls under this category makes use Weighted round robin method [13] [14].

▪ **Stateless adaptive**

These are those algorithms which take into consideration the server side information and are not concerned with current state of the client [13].

Fastest response time algorithm falls under this category of stateless adaptive algorithms..

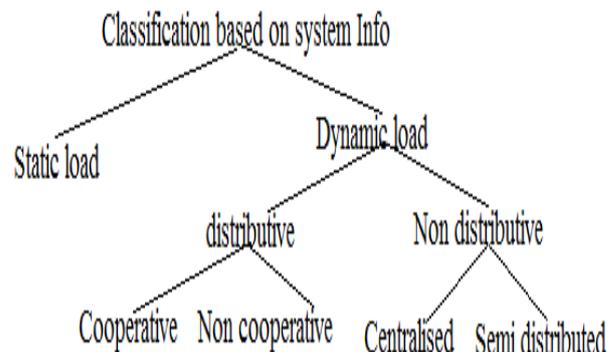
▪ **State full non adaptive**

These are those type of algorithms which take into account information pertaining to the client requests [13].

Algorithms such as weighted round robin algorithm, list based weighted round robin algorithm, Least connection-weighted least connections algorithm, Shortest expected delay, Never queue scheduling algorithm, Destination hashing locality based scheduling algorithm etc [13][21][22][23][24].

G. Classification based on the system information

Based on the system information required algorithms can be classified in to two types



- Static Load balancing algorithms
- Dynamic load balancing algorithms

Static load balancing algorithms

Algorithms which fall under this category require prior knowledge of the system and do not depend on the current state of the system [6]. Here, while balancing the load on the servers, the performance of the servers is determined and known prior to execution of new tasks [6]. With the information obtained from the previous tasks or before starting a new task, the load on the server is distributed based on the performance statistics obtained earlier. Here a master processor distributes the work and the slaves process estimate and calculate the load and send the results to the back to their master [6][8]. Keeping in mind to minimize the communication costs, the main goal of static load balancing algorithms is to reduce the execution times of the tasks [6].

Algorithms such as Round robin, randomized algorithm, Central manager algorithm, threshold algorithm etc fall under this category of static load balancing algorithms [6].

Dynamic load balancing Algorithms

Here, in dynamic load balancing algorithms, load balancing is done not based on prior information of the system but based on the current state of the system [6][7][12] The main difference between the static and dynamic algorithms is the calculation of load [6].

Central queue algorithm and local queue algorithm fall under this category of dynamic load balancing algorithms [6]. There are two kinds of dynamic load balancing algorithms:

- Distributed dynamic algorithms
- Non- distributed load balancing algorithms

Distributed dynamic algorithms

In the distributed algorithms, the execution and initiation of load balancing algorithm is carried out by all nodes connected and the resulting load which is calculated is shared and communicated by all the nodes in two ways[6], they are as follows:

- Co-operatively distributed
Here, in this setting, the nodes in a distributed mode work collectively and achieve objective goals [6].
- Non co- operatively distributed
In this type of distributed dynamic algorithms the nodes which are connected work individually to obtain objectively local goals [6].

Non-distributed dynamic algorithm

In the non distributed dynamic algorithm, not all nodes connected in a network or in system participate in the act of load balancing but only a single or a few nodes perform take up the responsibility of balancing the nodes [6]. The communication and sharing of load balance is done in two ways in non distributed algorithms, they are as follows:

- Centralized non-distributed setting
Here in this setting only a single node is responsible for balancing the load in system, all other nodes communicate with this single node [6].

- Semi-distributed setting

Here in this type of setting, nodes connected in a system are grouped into clusters and a single node in each cluster is responsible for the balancing of load, where the remaining clusters have to communicate with this central node in the cluster [6]. The overall load balancing is carried out by these collection of central nodes [6][11].

V. ANALYSIS AND DISCUSSION

An analysis made on the obtained results has led in identifying the benefits and shortcomings of scheduling algorithms. The advantage of Round robin algorithm is that it does not require much inter process communication but it has an drawback of not being able to achieve the expected levels of performance [6]. Similarly, the drawback of central manager algorithms is that it requires high levels of inter process communication which might create bottle neck problems [6].

VI. CONCLUSION

Through this report, different types of scheduling algorithms present for load balancing on web servers are thoroughly discussed, classified and evaluated. Also, benefits and shortcomings of these algorithms were identified. A complete classification and analysis of the different load balancing algorithms for web servers was discussed.

References

- [1] Abbas Karimi, Faraneh Zafarshan, Adzan b. Jantan, A.R Ramli, M.Iqbal b. saripan, "A new fuzzy approach for dynamic load balancing algorithm", *International Journal of Computer Science and Information security*, vol 6, no .1, 2009.
- [2] D. L. Eager, E. D. Lazowska, and J. Zahorjan, "Adaptive load sharing in homogeneous distributed systems," *IEEE Trans. Softw. Eng.*, vol. 12, pp. 662-675, 1986.
- [3] D. L. Eager, E. D. Lazowska, and J. Zahorjan, "A comparison of receiver-initiated and sender-initiated adaptive load sharing (extended abstract)," *SIGMETRICS Perform. Eval. Rev.*, vol. 13, pp. 1-3, 1985.
- [4] M. Livny and M. Melman, "Load balancing in homogeneous broadcast distributed systems," *in Proceedings of the Computer Network Performance Symposium*. College Park, Maryland, United States: ACM, 1982, pp. 47-55.
- [5] W. Leinberger, G. Karypis, and V. Kumar, "Load Balancing Across Near-Homogeneous Multi-Resource Servers," presented at Proceedings, *9th Heterogeneous Computing Workshop (HCW 2000)* Cancun, Mexico, 2000.
- [6] K.ramana, A subramanyam, A. Ananda rao, "Comparitive analysis of distributed webserver sstems load balancing using qualitative parameters", *VSRD-IJCSIT*, Vol. 1 (8), 2011, 592-600
- [7] S. Malik, "Dynamic Load Balancing in a Network of Workstation", 95.515 Research Report, 19 November, 2000.
- [8] Derek L. Eager, Edward D. Lazowska, John Zahorjan, "Adaptive load sharing in homogeneous distributed systems", *IEEE Transactions on Software Engineering*, v.12 n.5, p.662-675, May 1986.
- [9] G. R. Andrews, D. P. Dobkin, and P. J. Downey, "Distributed allocation with pools of servers," *in ACM SIGACT-SIGOPS Symp. Principles of Distributed Computing*, Aug. 1982, pp. 73-83

- [10] Zhong Xu, Rong Huang, "Performance Study of Load Balancing Algorithms in Distributed Web Server Systems", *CS213 Parallel and Distributed Processing Project Report*
- [11] Ali M. Alakeel, "A Guide to Dynamic Load Balancing in Distributed Computer Systems", *IJCSNS International Journal of Computer Science and Network Security*, VOL.10 No.6, June 2010
- [12] Y.Wang and R. Morris, "Load balancing in distributed systems," *IEEE Trans. Computing*, C-34, no. 3, pp. 204-217, Mar. 1985
- [13] S.kontogianis, S. Valsamidis, P. Eframidis, A.karakos, "An adaptive load balancing algorithm for cluster based web systems", <http://skontog.gr/papers/duthtr-12-07.pdf>
- [14] Batheja, J., and Parashar, M, "A framework for adaptive cluster computing using javaspaces", *Cluster Computing* vol. 6-3 (2003), 201-213.
- [15] Cardellini, V., Casalicchio, E., Colajanni, M., and Yu, P. S, "The state of the art in locally distributed web-server systems" *ACM Computing Surveys* vol. 34-2 (2002),263-311
- [16] Kant, K.; Won, Y, "Server capacity planning for web traffic workload" *IEEE Transactions on Data and Knowledge Engineering*, v.11, n.5, p.731-747.
- [17] Cardellini, V., Colajanni, M., and Yu, P. S. "Geographic load balancing for scalable distributed web systems", *In Proc. of 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (2000)*, pp. 20-28.
- [18] Casalicchio, E., and Colajanni, M, "A client-aware dispatching algorithm for web clusters providing multiple services", *WWW ACM (2001)*, 535-544.
- [19] Colajanni, M., Yu, P. S., and Dias, M. D, "Analysis of task assignment policies in scalable distributed web-server systems", *IEEE Trans. on Parallel Distributed Systems*, vol. 9-6 (1998), 585-597
- [20] Mamatas, L., and Tsaoussidis, V, "A new approach to service differentiation: Noncongestive queueing", *In Proc. of International Workshop on Convergence of Heterogeneous Wireless Networks (2005)*, pp. 78-83.
- [21] O'Rourke, P., and Keefe, M, "Performance Evaluation of Linux Virtual Server", *In Proc. of 15th System Administration Conference, LISA (2001)*, pp. 79-92.
- [22] Weinrib, A, and Shenker, S, "Greed is not enough: Adaptive load sharing in large heterogeneous systems", *In Proc. of IEEE INFOCOM '88 (1988)*, pp. 986-994.
- [23] Zhang, W, "Linux server clusters for scalable network services", *In Proc. of Ottawa Linux Symposium (2000)*, pp. 437-456.
- [24] Zhang, W, "Build highly-scalable and highly-available network services at low cost" *Linux Magazine (November 2003)*.
- [25] Teixeira, M, M. Santana, M. J.Santana, R. H. C, "Analysis of Task Scheduling Algorithms in Distributed Web-server Systems", *In International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS 2003)*, p.655-663. Montreal, Canada, jul., 2003.
- [26] Jiani Guo and Laxmi Narayan Bhuyan, "Load Balancing in a Cluster-Based Web Server for Multimedia Applications", *IEEE Trans. Parallel Distrib. Syst.* 17, 11 (November 2006), 1321-1334.

ABOUT THE AUTHOR



Sairam Vakkalanka is now pursuing masters in software Engineering at Blekinge Tekniska Högskola (Blekinge Institute of technology), Karlskrona, Sweden.

Design of Online Ups System with Over Voltage, Under Voltage and Phase Out Protection

Nagaraja Naik R,¹ Jayapal.R,²

^{1,2} (Department of E&EE R V College of Engineering Bangalore, Karnataka state, India)

Abstract: Uninterruptible power supplies (UPS) play an important role in interfacing critical loads such as computers, communication systems, medical/life support system and industrial controls to the utility power grid. They are designed to provide clean and continuous power to the load under essentially any normal or abnormal utility power condition. Among the various UPS topologies such as online UPS, offline UPS and line interactive UPS, online UPS are widely used. In this paper hardware implementation of single phase 50Hz, online uninterruptible power supply with over voltage, under voltage and phase out protection have been carried. Atmega32 microcontroller is the heart of the system and controls entire system. By programming the microcontroller using embedded C, SPWM pulses to drive H-bridge are generated. By alternating switching switches of two legs of H-bridge alternating 9V DC voltage is converted into 9V AC voltage. Output of H-bridge is given to step up transformer to step up the voltage to 220V, 50Hz. The microcontroller is so programmed that at every instant it checks the voltage that is supplied to the load through sensors. At any instant if it detects that there is over voltage, under voltage or phase out, microcontroller acts to isolate the load from the power source by sending tripping signal to relay. Once relay isolates the load from the power source, microcontroller supplies the load through the charged battery. Battery supplies the load until the power supply voltage reaches normal value of voltage.

Keywords: Online UPS, Offline UPS, Lineinteractive UPS SPWM, microcontroller ATMEGA32, Embedded C.

I. INTRODUCTION

An uninterruptible power supply (UPS), uninterruptible power source or sometimes called a battery backup is a device which maintains a continuous supply of electric power to connected equipment by supplying power from a separate source when utility power is not available. A UPS is inserted between the source of power (typically commercial utility power) and the load which is protected. When a power failure or abnormality occurs, the UPS will effectively switch from utility power to its own power source almost instantaneously [1]-[2]. While not limited to any particular type of equipment, a UPS is typically used to protect computers, telecommunication equipment or other electrical equipment where an unexpected power disruption could cause injuries, fatalities, serious business disruption or data loss [3]. UPS units come in sizes ranging from units which will back up a single computer without monitor (around 200 VA) to units which will power entire data centers or buildings (several megawatts). Larger UPS units typically work in conjunction with generators [1]-[8].

Conventional UPS topologies can mainly be categorized into three different types:

1. Off-line
2. line-interactive and

3. On-line [1].

In offline UPS under normal operation, a small amount of power is being converted from AC to DC to maintain battery charge. When input AC power goes out of specification, the inverter converts the DC power to AC to support the load. When the input power goes out of specification, there is a power disturbance in output voltage as the power failure is detected, the relay operates, and the output inverter turns on to begin to supply the load [1]. Though offline UPSs are very inexpensive and has high efficiency normal operation has following disadvantages [1]-[3].

- The offline UPS is normally only applied to single-phase (workstation-level) non-critical loads
- Its limitations, especially the generator incompatibility make it unsuitable for three-phase applications.
- Its application is limited for only low power applications

Line interactive resembles the offline UPS topology, but inserts a transformer or inductor in series between the utility power source and the load [10]-[12]. This inline inductor enables the UPS inverter to “interact” with incoming power and provide a measure of power conditioning to the load. This “buck-and-boost” circuitry helps with high and low input voltage conditions. Like the offline UPS, the line-interactive UPS can be inexpensive and efficient because they only support the entire critical load during power disturbances, and only for the duration of the battery. The line interactive UPS has following drawbacks [1]-[3]:

- Dynamic load changes cause power to be extracted from the battery. The resulting frequent hits on the battery can shorten battery life.
- Line-interactive UPS cannot completely isolate the critical load from the input line without operating on battery.
- Small perturbations in frequency and power quality can get passed directly to the critical load
- It can't be used for high power applications [6]-[9].

Among different types of UPS systems, the on-line UPS is the superior topology which not only overcomes the draw backs of the offline and line interactive ups but also has performance, power conditioning and load protection. Incoming AC power is rectified to DC power to charge battery of the UPS. The output inverter takes the DC power and produces regulated AC power to support the critical load. Battery is charged during normal operation. When the input power is out of specifications the batteries provide power to support the inverter and critical load.

Following are the advantages of online UPS

- The critical load is completely isolated from the incoming AC input power.
- The critical load is always being supplied by the output inverter, which is always being supplied from

the internal DC battery. When input power fails, there is no transitional sag in the output voltage because the inverter is already operating on DC input.

- The output inverter usually contains a step up or an isolation transformer. This enables the UPS to be electrically isolated and provide common mode noise protection for the load.
- A fault on the input line causes the UPS to go to battery power, but the UPS rectifier will not allow power from the DC battery to flow upstream [6], [8], [10] & [11].

II. PROPOSED BLOCK DIAGRAM

Block diagram of the proposed online UPS system with over voltage, under voltage and phase out protection using ATmega32 microcontroller is shown in Fig 1. The block diagram mainly consists of following important blocks:

ATmega32 microcontroller: It acts as the heart of the system. It controls and monitors entire system. The main function of this microcontroller is to generate SPWM signals. These signals are given to H-bridge switches to convert dc voltage to ac voltage. Microcontroller also takes care of the protection. It protects the load from over voltage, under voltage and phase out conditions by sending a tripping signal to relay. After relay isolates the load from supply the load is now supplied from battery unit. Due to the fluctuations of energy sources, which impose stringent requirements for inverter topologies and controls.

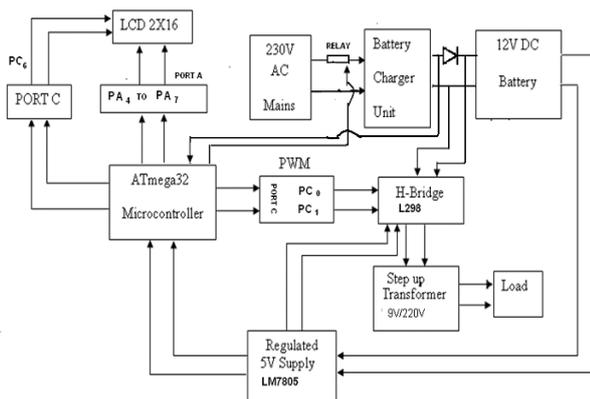


Fig 1: Block diagram of ONLINE UPS using AT-Mega32 microcontroller.

The function of an inverter is to change direct current (DC) input voltage to a symmetric alternating current (AC) output voltage of desired magnitude and frequency. When the main power is not available UPS uses batteries and inverter to supply AC power. A rectifier is used to recharge the battery used when the main power is back. Transformer is used to step up the voltage across the h-bridge to 220V [1]-[8].

In the present work design of online UPS system with over voltage, under voltage and phase out protection is taken up. The hardware prototype is implemented using Atmega32 microcontroller as control circuit. Microcontroller generates SPWM signals in order to drive H-bridge and to protect the load whenever there is change in voltage specifications through relay.

In the present work sinusoidal pulse width modulation (SPWM) technique is used to control the switches of the H-bridge. This technique is widely used in inverter to digitize the power so that a sequence of voltage pulses can be generated by the on and off of the power switches. The pulse width modulation inverter has been the main choice in power electronics, because of its circuit simplicity and rugged control scheme [9]. SPWM techniques are characterized by constant amplitude pulses with different duty cycle for each period. The width of this pulses are modulated to obtain inverter output voltage control and to reduce its harmonic content. In the present work is to replace the conventional method with the use of ATmega32 microcontroller. It is also low cost and has a small size of control circuit for the single phase H-bridge inverter [9]-[12].

III. FLOW CHARTS AND DESIGN DETAILS

ATmega32 can operate at a maximum frequency of 16MHz. In the present work 1 MHz frequency is selected. Timer/counter control register (TCCR0) is an 8 bit register. The register description of TCCR0 is as shown in Fig 2.

Bit	7	6	5	4	3	2	1	0	
	FOC0	WGM00	COM01	COM00	WGM01	CS02	CS01	CS00	TCCR0
Read/Write	W	R/W	R/W	R/W	R/W	R/W	R/W	R/W	
Initial Value	0	0	0	0	0	0	0	0	

Fig 2: Register description of TCCR0

In the present work only bit 1 (CS01) is used. For TCCR0=2, CS01 (clock select) is, $F_C/8 = 1\text{MHz}/8 = 125\text{KHz}$

Total time period is calculated by eqn. 1.

$$T = \frac{1}{f} = \frac{1}{125\text{KHz}} = 8\mu\text{s} \quad (1)$$

To get 50Hz output time period is given from eqn. 1.

$$T = \frac{1}{50} = 20\text{ms}$$

50% Duty cycle is selected, duty cycle is given by eqn. 2.

$$D = \frac{T_{ON}}{T} \quad (2)$$

where, T_{ON} = ON time

$$0.5 = \frac{T_{ON}}{20\text{ms}}, \quad T_{ON} = 10\text{ms} \ \& \ T_{OFF} = 10\text{ms} \quad (3)$$

10ms time period is decomposed to reduce the harmonic content. The graphical view of switching pulses is as shown in Fig 3.

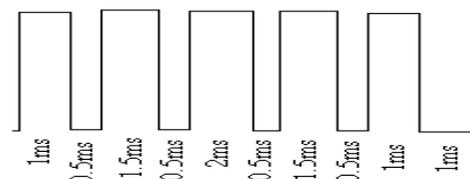


Fig 3: Graphical view of switching pulses.

To load 1ms, 0.5ms, 1.5ms in program it should be divided by $8\mu\text{s}$ we get,

$$1\text{ms} = \frac{1000\mu\text{s}}{8\mu\text{s}} = 125; 0.5\text{ms} = \frac{500\mu\text{s}}{8\mu\text{s}} = 62.5; 1.5\text{ms} = \frac{1500\mu\text{s}}{8\mu\text{s}} = 187.5 \quad (4)$$

For 8 bit timer maximum value is 255(decimal). To load 125, 62 and 187 into timer/counter (TCNT0) it should be subtracted from 255. The obtained values are preloaded to get sinusoidal pulse width modulation. The dead time for the PC2 and PC3 switching pulse are loaded.

The dead time is set as 1ms which is acceptable for various types of power transistor. Typically, the switching devices consume only a few nanoseconds to operate as a switch. This delay time is necessary to avoid the damage on the inverter circuit during the switching pair transition.

The flowchart explaining sinusoidal pulse width modulation signal generation is as shown in Fig 4. The flowchart explaining the over voltage, under voltage and phase out protection is shown in Fig 5.

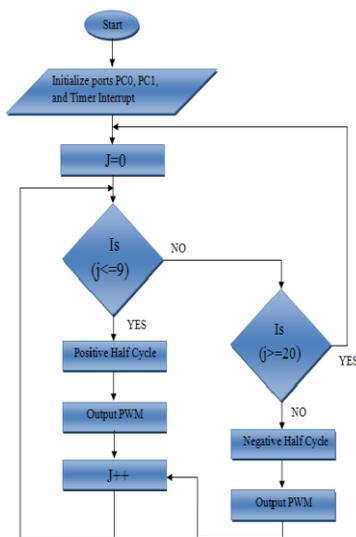


Fig 4: Flow chart for SPWM generation

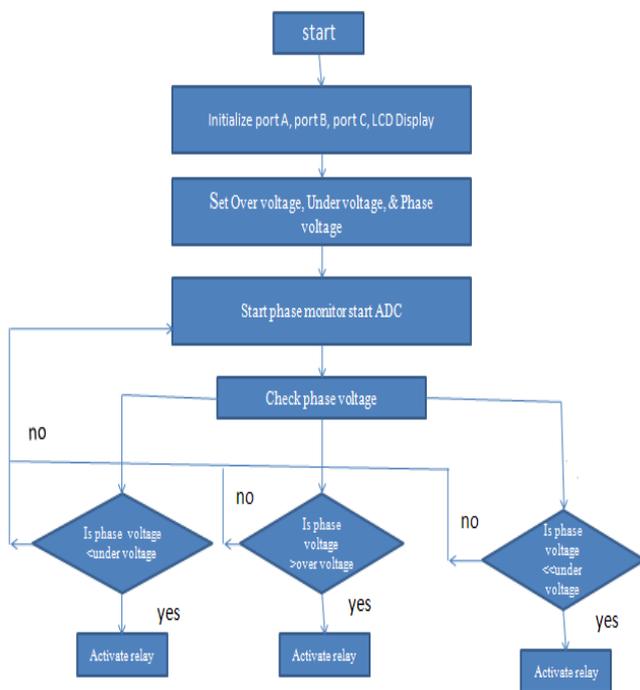


Fig 5: Flow chart for over voltage, under voltage and phase out protection

IV. INVERTER DESIGN

Design of 50 Watt Inverter, assuming efficiency of the inverter to be 98%

$$\text{Efficiency} = P_o / P_{in}; \quad (4)$$

$$P_{in} = P_o / 0.98;$$

$$= 50 / 0.98;$$

$$P_{in} = 51 \text{ Watts};$$

Where,

$$P_{in} = \text{DC input power to the inverter} = V_{dc} * I_{dc};$$

P_o = AC output power of the inverter

Since the input DC voltage (V_{dc}) is varied from 12.5-14V, the range of input DC current (I_{dc}) is 4A to 3.64A.

Assuming the output power factor of the inverter to be 0.8; Therefore the output power of the inverter

$$(P_o) = V_{rms} * I_{rms} * \cos\phi; \quad (5)$$

Since the output voltage of the inverter (V_{rms}) = 220V;

The output current of the inverter (I_{rms}) = 0.284Amps.

V. Experimental Results And Analysis

Complete fabricated hardware set up is as shown in Fig 7.

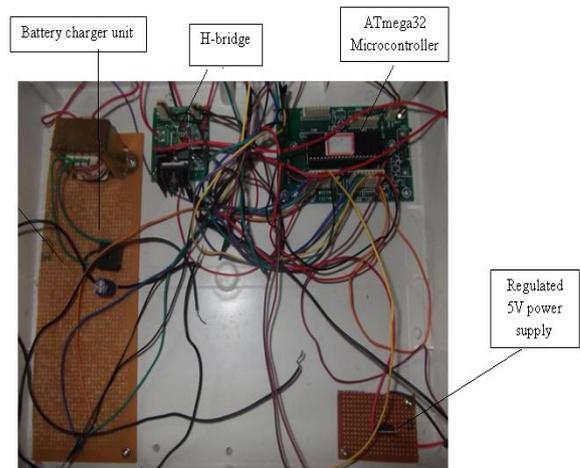


Fig 7: Complete fabricated hardware set up

The output waveform across two legs of the H-Bridge is shown in Fig 8. The out of H-bridge is given to the step-up transformer. The output voltage of step up transformer is 219V. The PWM output obtained across H - Bridge was of 50Hz with train of pulses switched at a frequency of 500Hz.

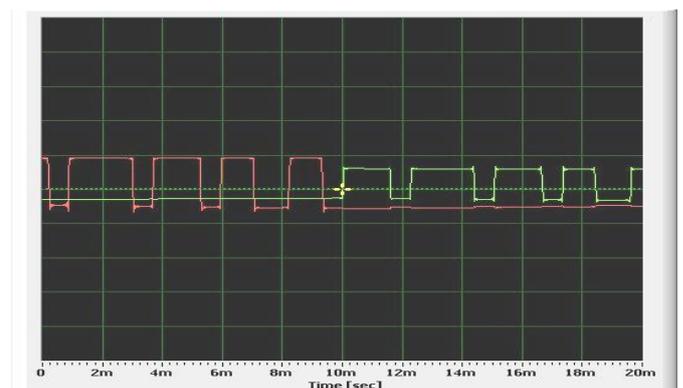


Fig 8: Output waveform across the two legs of the H-bridge

Fig 9 shows the experimental setup of project work with 15W compact fluorescent lamp (CFL) connected across step-up transformer secondary as a load. The waveform for the same is shown in Fig 10 from the waveform it can be seen the even under load the PWM output across H-bridge was almost similar to that of No load condition.

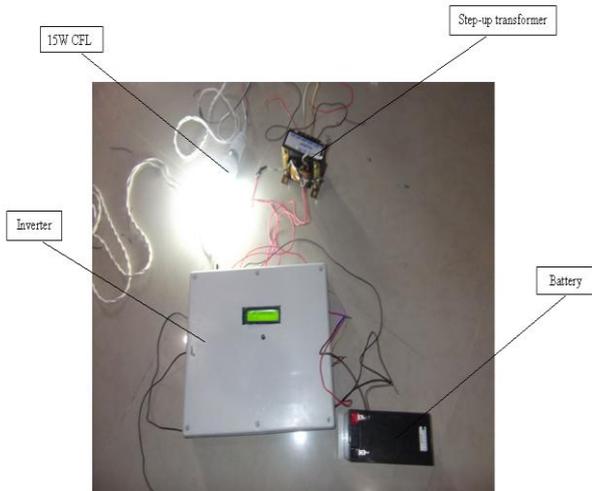


Fig 9: Snapshot of complete hardware with 15W CFL as the load

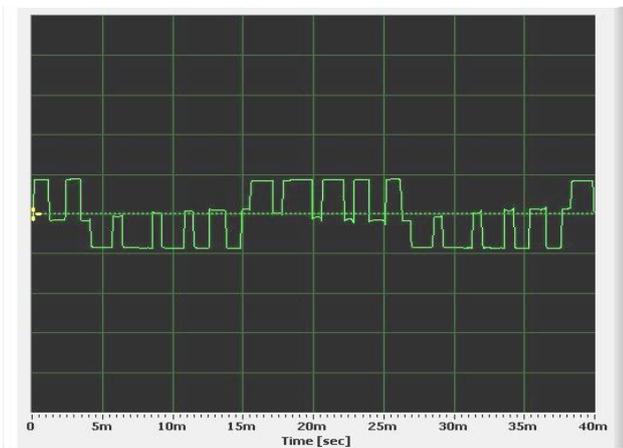


Fig 10: PWM output across H-bridge with load of 15W CFL

The various readings of CFL with variation in load is tabulated in Table 1. From this table it can be observed that the designed online UPS works efficiently well under full load condition.

Table 1: Readings of CFL as a load

Sl.No	CFL(W)	V _{dc} (V)	I _{dc} (A)	P _{dc} (W)	V _{rms} (V)	I _{rms} (A)	P _{ac} (W)	%η
1	15	12.6	2.8	35.28	218	0.11	23.98	67.97
2	30	12.5	3.2	40	216	0.14	30.24	75.6
3	40	12.4	3.5	43.4	213	0.18	38.34	88.34
4	45	12.00	3.6	44	210	0.2	42	93.33

Plot of power versus efficiency is as shown in Fig 11, from which it can be observed that as the load increases the efficiency increases.

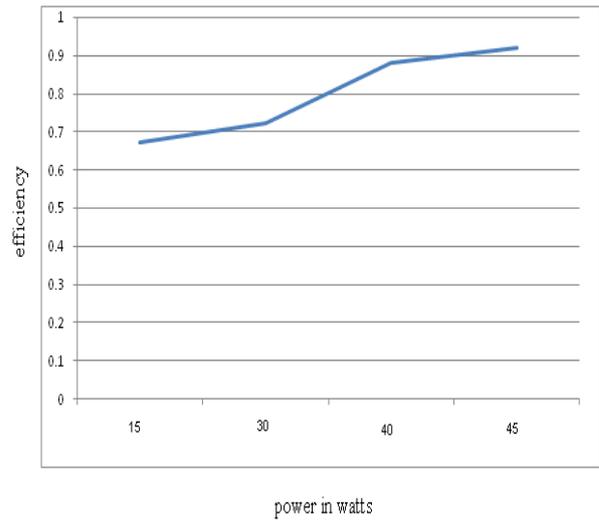


Fig 11: Plot of power versus efficiency

The hardware was also tested for over voltage by setting the nominal voltage of 180V. On supplying a voltage of 220V to the system, the microcontroller senses the over voltage through sensors and sends the tripping signal to relay and correspondingly relay acts and isolates the load from the supply. The load is now supplied by battery until the supply voltage comes close (i.e.±2% tolerance) to nominal voltage which is set in the microcontroller.

The Hardware is tested for phase out and under voltage by setting the nominal voltage of 240V. When a voltage of 220V is applied to the system, the microcontroller sense the under voltage through sensors and sends the tripping signal to relay. Relay acts correspondingly and isolates the load from the supply. The load is supplied by battery until the supply voltage comes close to nominal value set in the microcontroller.

VI. CONCLUSION

In the present work, online ups system with over voltage under voltage and phase out protection has been implemented using ATmega32 microcontroller and H-bridge.

Important conclusions that are drawn out of the investigations in the present work are:

- Output waveforms of the UPS with and without load were found to be satisfactory and were in accordance with the design.
- Pulse width modulation (PWM) circuit is implemented in a single board ATmega32 microcontroller, which makes system reliable, compact.
- In addition, with the high programming flexibility, the design of the switching pulse can be further altered easily without any further changes on the hardware.
- H-bridge based on L298 integrated circuit is used which gives better efficiency and makes the system compact.
- Working of online ups with power supply on and with power supply off where found out to be satisfactory.
- The hardware designed isolates the load from the supply in case of over voltages, under voltages and phase outs. The battery supplies the voltage supply under those conditions. Thus supplying continuous supply to the load.

REFERENCES

- [1]. S. Martínez, M. Castro, R. Antoranz, and F. Aldana, "Off-line uninterruptible power supply with zero transfer time using integrated magnetics," *IEEE Trans. Ind. Electron.*, vol. 36, no. 3, pp. 441–445, Aug. 1989.
- [2]. M. T. Tsai and C. H. Liu, "Design and implementation of a cost-effective quasi line-interactive UPS with novel topology," *IEEE Trans. Power Electron.*, vol. 18, no. 4, pp. 1002–1011, Jul. 2003.
- [3]. Adel Nasiri, Zhong Nie, Stoyan B. Bekiarov, and Ali Emadi, "An On-Line UPS System With Power Factor Correction and Electric Isolation Using BIFRED Converter" *IEEE transaction on Industrial Electronics*, vol. 55, no. 2, pp. 722 - 730, February 2008.
- [4]. Kai Zhang, Yong Kang, Jian Xiong, and Jian Chen, "Direct Repetitive Control of SPWM Inverter for UPS Purpose" *IEEE Transaction on Power Electronics*, vol. 18, no. 3, pp. 784-792, may 2003.
- [5]. T. Kawabata, T. Miyashita, N. Sashida, and Y. Yamamoto, "Three phase parallel processing UPS using multi-functional inverter," *Proc. IEEE Industry Applications Soc. Annu. Meeting Conf.*, Oct. 1989, vol. 1, pp. 982–987.
- [6]. J. C. Wu and H. L. Jou, "A new UPS scheme provides harmonic suppression and input power factor correction," *IEEE Trans. Ind. Electron.*, vol. 42, no. 6, pp. 629–635, Dec. 1995.
- [7]. M. Pacheco, L. C. de Freitas, J. B. Vieira, Jr., A. A. Pereira, E. A. A. Coelho, and V. J. Farias, "An online no-break with power factor correction and output voltage stabilization," *IEEE Trans. Power Electron.*, vol. 20, no. 5, pp. 1109–1117, Sep. 2005.
- [8]. G. Joos, Y. Lin, P. D. Ziogas, and J. F. Lindsay, "An online UPS with improved input-output characteristics," *Proc. IEEE Applied Power Electronics Conf. and Expo.*, Feb. 1992, pp. 598–605.
- [9]. Y. Lin, G. Joos, and J. F. Lindsay, "Performance analysis of parallel processing UPS systems," in *Proc. IEEE Applied Power Electronics Conf. and Expo.*, Mar. 1993, pp. 533–539.
- [10]. S. Rathmann and H. A. Warner, "New generation UPS technology, the delta conversion principle," *Conf. Proc. 31st IEEE Industry Applications Soc. Annu. Meeting*, Oct. 1996, vol. 4, pp. 2389–2395.
- [11]. F. Kamran and T. G. Habetler, "A novel on-line UPS with universal filtering capabilities," *IEEE Trans. Power Electron.*, vol. 13, no. 3, pp. 410–418, May 1998.
- [12]. S. J. Jeon and G. H. Cho, "A series-parallel compensated uninterruptible power supply with sinusoidal input current and sinusoidal output voltage," in *Proc. 28th IEEE Power Electronics Specialists Conf.*, Jun. 1997, vol. 1, pp. 297–303.

Method of Performance Evaluation of WANET Using NS-2

Ms.Dharmistha D. Vishwakarma

Department of Elec. Engg., Faculty Tech. & Engg., M.S.University of Baroda, Vadodara

ABSTRACT: Next Generation Network demands to exchanging information at anytime using wireless network with a high data rate. To satisfy the necessity of today's generation the network requires which can be establish anywhere anytime. Wireless ad hoc Network (WANET) can established anytime anywhere without any backbone infrastructure using the mobile nodes. WANET uses wireless mobile nodes that operate as terminals as well as routers in the network, without any centralized administration. The performance of WANET can be measure using the simulation in the NS-2 tool. The NS-2 is a widely used simulator for the performance evaluation of WANET. This paper describes that the creation of wireless network and its performance evaluation using the awk script.

Keywords: AODV, NS-2, Packet Delivery Ratio, Trace file, , WANET

I. INTRODUCTION

Ad-Hoc networking is a concept in wireless networking, which means that users wanting to communicate with each other form a temporary network, without any form of centralized administration. Each node participating in the network acts both as host and a router and therefore is willing to forward packets for other nodes, which makes use of a routing protocol. Each node has to maintain some form of information regarding the network around it, and some algorithm governing the sending and receiving of data packets. This algorithm, together with the supporting information regarding network conditions, is called a Routing Protocol [1, 2].

Characteristics of Ad hoc networks impose a set of new demands on the routing protocol. The most important characteristic is the dynamic topology, which is a due to the node mobility, nodes can change position quite frequently, which means that there is a need a routing protocol that quickly adapts to topology changes. Due to power limitation, the routing protocol should try to minimize control traffic, such as periodic update messages, conserving memory, power and bandwidth resources. The main function of the routing protocol is to detect and maintain the optimal route to send data packets between a source and destination via intermediate node(s).

NS-2 is the discrete event network simulator initially developed for the wired technology in 1989 with the development of the real network simulator. Monarch group from the Dept. of computer science at the University of Rice has done the necessary extension to the NS-2 for the simulation of wireless and mobile network. NS-2 uses an uncommon programming technique called split-level programming [3, 4]. In NS-2 C++ is used to provide fast computable parts of network nodes that are glued together by the tcl scripting language. Due to the combination of C++ and tcl script it provides fast computation and flexibility to change the node structures. Section II explains

the scenario of the different simulation tools and structure of Network Simulator tool and the method to create the tcl file for the wireless network section III describes that how to evaluate performance of the any wireless network using the awk script, section IV concludes the paper for the evaluation of performance of wireless network using NS-2.

II. SIMULATION TOOLS

The physical world is modeled in wireless network simulation software to evaluate the behavior and performance of protocols for wireless networks. It is a good compromise between cost, complexity and accuracy of the results. Models can capture reality only to a limited extent. Simulation results will generally not be as accurate as real implementations. There are many simulation tools are available for the simulation of wireless networks or the next generation network.

GloMoSim is developed at UCLA (California, USA). It is a scalable and parallel simulator for wireless networks written in PARSEC (PARallel Simulation Environment for Complex Systems). It consists of a set of library modules which were developed using PARSEC. PARSEC is a C-based parallel simulation language for evaluating a variety of wireless network protocols. It is an open-source network simulator which is based on parallel programming. Glomosim supports TCP Family, IEEE 802.11 CSMA/CA, MAC, UDP, CBR, FTP, HTTP, Fisheye, LAR, ODMRP, WRP, DSR, MACA, Telnet, AODV etc. protocols, wireless. It can emulate the real world more accurately. Easy configurable and fast simulation Many protocols already implemented, the help is very well documented, Dynamic behaviour of the network can be visualize using VT tool. GloMoSim currently solely supports wireless networks, it provides more physical-layer models than ns-2. It may be hard to debug the parallel programs [5].

Qualnet is the commercial network simulation tool not freely available to the user. It is extension of the GLOMOSIM. It is GUI based simulator and easy to create the simulation wireless network. QualNet can support real-time and faster than real-time simulation speed. [6].

OMNet++ is a well-designed simulation package written in C++. OMNET++ is capable of simulating any system composed of devices interacting with each others. The basic entity in OMNET++ is a module. Modules can be composed of submodules or they can be atomic. Only atomic modules capture the actual behavior. Modules communicate with each other via messages through gates. Gates are linked to each other using connections. A connection can be associated with a propagation delay, error rate and data rate. OMNET++ does not have an energy model [7].

OPNET was first proposed by MIT in 1986. It is first commercial available simulation tool & now the most widely used GUI based commercial network simulator. OPNET has features of: Modeling & Simulation cycle,

hierarchical modeling, specialized for communication networks OPNET Consists of Three layers: The Network Model, The Node model, The Process model OPNET provides a flexible, high-level programming language with extensive support. It Supports Routing protocols OSPF, RIP, EIGRP, BGP, IGRP, DSR, TORA IS-IS, PNNI, MAC, mobility of nodes, ad hoc connectivity, different application models, Node failure models, modeling of power-consumption etc. Very well documented help is provided. This tool has Relatively high price, complex, takes time to learn, is restriction to its portability [9].

In NS-2 conventional mobile nodes consist of a number of C++ objects which are bond together by tcl-scripts (split-level-programming). Every box (square, ellipse, or trapezoid) is a C++ object. All objects are bond together via tcl-commands. Routing within the mobile gateway is carried out by the new object AGW Agent.It forwards data packets to the inner or outer wireless interface respectively. Two modified routing agents (RTAgent) then process all data and routing packets, in order to perform reasonable ad-hoc behavior [10]. NS-2 supports TCP family, UDP, CBR, FTP, HTTP, Pareto, Exponential protocols, wires, wireless, unicast, multicast routing types.

As a first approximation, the ns-2 could seem a bit difficult to work with. In order to ease its utilization,its structure is explained next. The ns-2 employs two languages. Firstly, the main characteristics of the scenario to simulate are described by OTcl [11, 12]. Secondly, the kernel of ns-2 is specified by C++. It is organized in different folders; each one corresponds to a specific protocol. For example, there is a folder called AODV where all the files that are necessary for the implementation of this protocol are stored. If a user wants to utilize the already developed software that is included in the ns-2 release, he/she will have to know how to specify the scenario by OTcl. However, if some new protocols want to be tested or studied, the inclusion requires the modification of C++ files. In our experience, the modifications are not too troublesome but finding where to modify is usually quite hard. Following are the steps to simulate the wireless network in NS-2 and to creating the tcl file.

- Create an instance of the simulator:
- Setup trace support by opening
- Next create a topology object
- Create the object God, "God (General Operations Director)
- Configuring the nodes
- Create nodes and the random-motion for nodes is disabled
- Give nodes positions to start with
- Setup node movement as the following example
- Setup traffic flow between the two nodes as follows: TCP connections between nodes
- Define stop time when the simulation ends and tell nodes to reset which actually resets their internal network components.
- Finally the generated file for any particular protocol is run and the output files are generated namely out.tr and out.nam and filename.tcl.

To run the simulation type on command prompt in terminal
 \$ ns filename.tcl
 \$ nam out.nam

Fig 1 and Fig. 2 shows scenario to visualize the behavior of the wireless ad-hoc network with the wireless network that exchange the packet and shows the coverage of each node in.

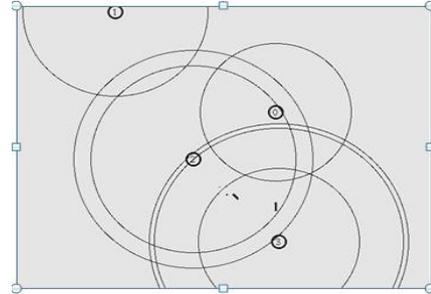


Fig 1: nam output of the wireless network with 4 nodes

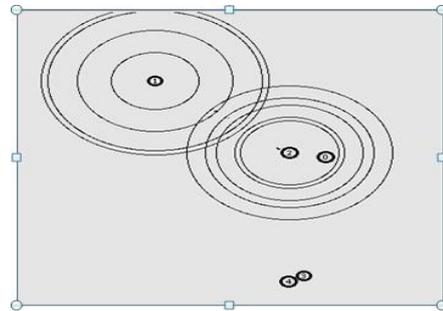


Fig 2: nam output of the wireless network with 5 nodes

III. PERFORMANCE EVALUATION USING NS-2

After finishing of writing tcl file for the network we can visualize the packet communication between the nodes participating in the network when the tcl file is executed. It also generate the trace file which contains the all the data related to packet communication between the nodes. In “*.tr” file each line consists of following information as shown in Fig 3.

Event	Time	From node	To node	Packet type	Packet Size	Flags	Fid	Src addr	Dest addr	Seq number	Pkt id
r:	1.3556	3	2	ack	40	-----	1	3.0	0.0	15	201
+	1.3566	2	0	ack	40	-----	1	3.0	0.0	15	201
-	1.3566	2	0	ack	40	-----	1	3.0	0.0	15	201
r	1.35576	0	2	tcp	1000	-----	1	0.0	3.0	29	199
+	1.35576	2	3	tcp	1000	-----	1	0.0	3.0	29	199
d	1.35576	2	3	tcp	1000	-----	1	0.0	3.0	29	199
+	1.356	1	2	che	1000	-----	2	1.0	31	157	207
-	1.356	1	2	che	1000	-----	2	1.0	31	157	207

Fig.3: Format of Trace file

Each line consists of:

- Event Descriptor (+, -, d, r)
- Simulation time (in seconds) of that event
- From Node & To Node, which identify the link on which the event occurred
- Packet type
- Packet size

- Flags (appeared as "-----" since no flag is set). Currently, NS implements only the Explicit Congestion Notification (ECN) bit, and the remaining bits are not used.
- Flow id (fid)
- Source and destination address in forms of "node.port".
- The network layer protocol's packet sequence number. What about UDP?
- The last field shows the unique id of the packet.

Each trace line starts with an event (+, -, d, r) descriptor followed by the simulation time (in seconds) of that event, and from and to node, which identify the link on which the event occurred. The next information in the line before flags (appeared as "-----" since no flag is set) is packet type and size (in Bytes). Currently, NS implements only the Explicit Congestion Notification (ECN) bit, and the remaining bits are not used. The next field is flow id (fid) of IPv6 that a user can set for each flow at the input OTcl script. Even though fid field may not use in a simulation, users can use this field for analysis purposes. The fid field is also used when specifying stream color for the NAM display. The next two fields are source and destination address in forms of "node.port". The next field shows the network layer protocol's packet sequence number. Note that even though UDP implementations do not use sequence number, NS keeps track of UDP packet sequence number for analysis purposes. The last field shows the unique id of the packet.

We need to extract the necessary data out of this file. GREP is a UNIX command to filter a file. The name comes from "search Globally for Lines Matching the Regular Expression and Print them". It takes a regular expression on the command line, reads the standard input or list of files, and outputs the lines containing matches for the regular expression.

Another option is to use AWK file. AWK is a general purpose computer language. AWK is designed for processing text-based data, either in files or data streams. The name AWK is derived from the surnames of its authors Alfred Aho, Peter Weinberger, and Brian Kernighan.

To run AWK file following commands are needed:

```
awk -f file1.awk file2.txt
awk -f file1.awk file2.txt > out.txt
```

Where,

```
file1.awk: is a command file
file2.txt: is a primary input file
out.txt: is an output file
```

A typical AWK program consists of a series of lines; each of them is on the form

```
/pattern/ {action}
```

Pattern is a regular expression Action is a command.

- Most implementations of AWK use extended regular expressions by default.
- AWK looks through the input file; when it finds a line that matches pattern, it executes the command(s) specified in action.
- awk or perl scripting language is used to interpret the trace file, it extract data from trace file trace.tr.
- Write the awk script (that is similar to C language) and save it in *****.awk:

Since perl script is very efficient in match data file which has the form of line records [6, 7]. And is very convenient in extract numeric variables from strings, also it is easy to do the computation as C programming language does.

Following is the program segment of awk file written for calculating packet delivery ratio for the trace file for wireless network example.

```
# awk file to find packet delivery ratio.
for (packet_id = 0; packet_id <= highest_packet_id;
packet_id++){
if ((send_time[packet_id]!=0) &&
(rcv_time[packet_id]!=0)){
start = send_time[packet_id];
end = rcv_time[packet_id];
packet_duration = end-start;
}
else
packet_duration = -1;
if (packet_duration > 0)
{packet_no++;
total_delay = total_delay + packet_duration;
}
printf("%d %f %f\n", packet_no, total_delay,
total_delay/packet_no);
}
```

After applying the awk file to the network user can note down the result for different situation of the network and then it will be converted to the graphical format using XGRAPH.

IV. CONCLUSION

The paper describes the different network simulation tools and their special features with the supporting facility of the WANET. It also explains the method for evaluation mechanism developed in NS-2 is really effective for scalable performance test in NS-2. It also could be easy to use for measure the network routing protocols' performance, meanwhile, since it has the fix model of analysis of the trace file, with some minor modification, it will then be apply to measure other kinds of evaluation with the whole network simulation. Since here only some important point and fields are discussed to for the trace file and awk file. However the evaluation scheme is very much tedious job than the simulation of the network.

V. ACKNOWLEDGEMENTS

The project work was carried out at Dept. of Elect. Engg, Faculty of Tech. & Engg. M.S.University of Baroda, Vadodara, Guj, India. Author is thankful to Department of Elect. Engg. for technical help in doing this work.

REFERENCES

- [1] S.-C. Woo, S. Singh, Scalable routing protocol for ad hoc networks, *Wireless Networks* 7 (5), 2001,513-529.
- [2] Dr.Yogesh Chaba and Naresh Kumar Medishetti, Routing Protocols in mobile Ad hoc Networks- A simulation Study, *Journal of Computer Science, JCS Vol.1 No.1, August 2005,pp 83-88.*
- [3] The Network Simulator - ns-2, <http://www.isi.edu/nsnam/ns/>, last visited may 26th, 2004.
- [4] "The CMU Monarch Project's Wireless and Mobility Extensions to NS. <http://www.monarch.cs.cmu.edu/>,"
- [5] L. Bajaj, M. Takai, R. Ahuja, R. Bagrodia, and M. Gerla. GloMoSim: A scalable network simulation environment. Technical Report 990027, UCLA Computer Science Department, 13, 1999

- [6] Scalable Network Technologies (SNT). QualNet.
<http://www.qualnet.com>
- [7] Drytkiewicz, W., S. Sroka, V. Handziski, A. Koepke and H. Karl, A mobility framework for omnet++, in: 3rd International OMNeT++ Workshop, at Budapest University of Technology and Economics, Department of Telecommunications Budapest, Hungary, 2003
- [8] G. F. Lucio, M. Paredes-Farrera, E. Jammeh, M. Fleury, and M. J. Reed. Opnet modeler and ns-2 - comparing the accuracy of network simulators for packet-level analysis using a network testbed. WSEAS Transactions on Computers, 2(3):700–707, July 2003
- [9] Alex Ali Hamidian, A Study of Internet Connectivity for Mobile Ad Hoc Networks in NS 2, Department of Communication Systems, Lund Institute of Technology, Lund University, January 2003
- [10] Introduction to PerlScript.
<http://www.cpan.org/authors/id/M/MS/MSERGEANT/PSIntro.html>
- [11] PerlTutorial. <http://www.comptechdoc.org/independent/perlmanual>

An Extensive Input Voltage and Fixed-Frequency Single Stage Series-Parallel LLC Resonant Converter for Dc Drive

P.Ganesh¹, T.Manokaran²

^{1,2}Department of Electrical and Electronics Engineering, Sri Subramanya College of Engineering and Technology
Anna university, Chennai.India

ABSTRACT: In this paper design of a single-stage LLC resonant converter are presented. A single-stage converter uses only one control signal to drive two power converters, a power factor corrector (PFC) converter and a dc/dc converter, for reducing the cost of the system. However, this simplicity induces power imbalance between two converters, and then, the bus voltage between two converters drifts and becomes unpredictable. To ensure that the bus capacitor voltage can be kept in a tolerable region, the characteristics of a PFC converter and an LLC tank are investigated, and then, a design procedure is proposed correspondingly. Finally, a single-stage LLC resonant converter is implemented to verify the analysis.

Key words: AC-DC power conversion, resonant power conversion.

I. INTRODUCTION

A conventional power supply was designed with two cells: The first cell functions as a Power Factor Corrector (PFC), and the second cell is a dc/dc converter which regulates system output voltage. To reduce the cost and complexity of the power supply, a single-stage topology [1]–[10] which uses only one control signal to drive two converters is presented.

It is common to insert a bulk capacitor between a PFC converter and a dc/dc converter to eliminate low-frequency noise at the system output [6]. However, since there is only one control signal in a single-stage converter, the power passing through the PFC converter is not equal to the power passing through the dc/dc converter. Therefore, the capacitor voltage drifts and becomes unpredictable. An LLC resonant converter is employed as the dc/dc converter because LLC resonant converters have characteristics of high efficiency and low noise [10]–[13].

The works in [6]–[9] make efforts to control the drifting bus voltage. However, since lower cost is a major advantage in a single-stage converter, keeping the system simple and at low cost is important. The work in [10] investigated the characteristics of a single-stage LLC converter. However, the LLC converter characteristics are not well derived, so the bus voltage cannot be limited as much as well. In this paper, the relationship between the output powers of a boost stage and an LLC stage is derived.

Then, a single-stage LLC converter design procedure is proposed, which ensures that the bus capacitor voltage can be kept in a tolerable region. Finally, an experimental circuit is implemented to verify the analysis.

In resonant topologies, Series Resonant Converter (SRC), Parallel Resonant Converter (PRC) and Series Parallel Resonant Converter (SPRC, also called LCC resonant converter) are the three most popular topologies. The analysis and design of these topologies have been

studied thoroughly. In next part, these three topologies will be investigated for front-end application.

1.1. CIRCUIT DESCRIPTION

Fig.1.1.1 shows a single-stage ac/dc LLC resonant converter. There is a bus capacitor between two converters, a boost like converter and a typical LLC converter. Two switches Q1 and Q2 are used to control these two converters.

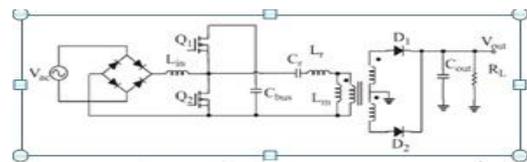


Fig.1.1.1: Single Stage LLC Converter

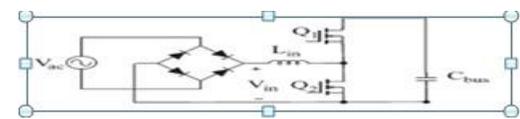


Fig.1.1.2: Simplified Circuit of the First Stage

The LLC converter is typical that has been presented in [7]–[10]. In the following, the operation details of the boost converter are introduced. Fig.1.1.2 shows the simplified circuit of the boost stage. The left-hand side of input inductor L_{in} is a rectified half-wave voltage source. The input voltage source charges input inductor L_{in} when the switch turns on. The energy stored in L_{in} is released to bus capacitor C_{bus} when the switch turns on. Because the boost cell uses the same switches with those in the LLC cell, the duty cycles of Q1 and Q2 are kept 50%.

Notice that the output power of this stage increases while the operating frequency decreases, because the inductor stores less energy in high frequency. The trend is the same with the LLC converter. Therefore, the fact that two converters can use the same control signal can be believed.

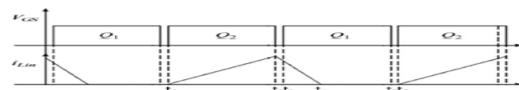


Fig.1.1.3 : Timing Diagram of Switch and inductor

The operational waveforms of the first stage are shown in Fig. 1.1.3.

Stage I— $t1-t2$: Switch Q2 turns on at $t1$. The voltage across the inductor is V_{in} , and then, the inductor is charged by the slope

$$\frac{di_{L_{in}}}{dt} = \frac{V_{in}}{L_{in}} \quad \dots (1)$$

Stage II— t_2-t_3 : Switches Q1 and Q2 are turned off in this stage. The inductor current flows into the body diode of Q1 that helps achieve zero-voltage switching (ZVS).

Stage III— t_3-t_4 : Switch Q1 turns on. The inductor current falls by the same slope

$$\frac{di_{Lin}}{dt} = -\frac{V_{cbus} - V_{in}}{L_{in}} \dots (2)$$

To reduce the higher frequency harmonics of the inductor current, the system operates in discontinuous-conduction mode (DCM). Therefore, the period of falling slope has to be less than the one of rising slope. That means that the bus capacitor voltage should be twice larger than the input peak voltage as follows:

$$V_{cbus} > 2V_{in,max} \dots (3)$$

Stage IV— t_4-t_5 : There is no energy left in the input inductor. The inductor current is kept zero.

Stage V— t_5-t_6 : Switches Q1 and Q2 turn off. Because the input impedance of the LLC converter is inductive, the input current of the LLC converter is lagged by the input voltage of the LLC converter. Therefore, the body diode of Q2 can be recharged by the LLC input current to achieve ZVS.

II. BUS VOLTAGE ESTIMATION

2.1 ENERGY EQUILIBRIUM

How to keep the voltage of the bus capacitor in a reasonable region is a major consideration in a single-stage converter analysis. Because there is only one control signal for two converters, the bus capacitor voltage varies due to energy imbalance of the two converters.

The bus capacitor voltage can be built by energy balance of two converters as shown in Fig. 2.1. When the boost output power P_{boost} is larger than the LLC output power P_{LLC} , the bus voltage will rise until the two power flows are equal. Therefore, theoretically, given a switching frequency f_{sw} , the steady-state bus voltage V_{bus} can be found.

In the following, the relationship between the output power and the operating frequency of the boost stage and the LLC stage is derived.

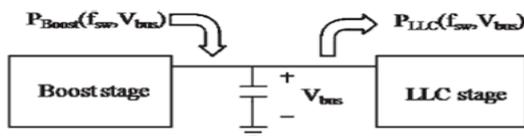


Fig 2.1.1 : Energy Balance At The Bus Capacitor

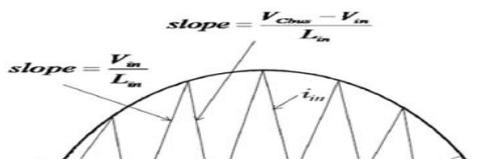


Fig. 2.1.2 : Waveform of the Input Voltage and Current

2.2 BOOST STAGE

First, consider the boost stage converter. Assume that the bus capacitor is large enough so that the voltage across the bus capacitor is stable in one cycle of an ac source. Given the inductor size L_{in} , the designed bus

capacitor voltage V_{bus} , and the ac source voltage V_{ac} , the waveforms of the input current and voltage can be determined as in Fig. 2.1.2. The average input current can be derived from (1) and (2)

$$i_{avg}(t) = \frac{V_{in}(t)T_{sw}}{8L_{in}} \left(\frac{V_{CBUS}}{V_{cbus} - V_{in}(t)} \right) \dots (4)$$

$$P_{boost} = \frac{\int_0^{T_{sw}} i_{avg}(t)V_{in}(t) dt}{T_{sw}} \dots (5)$$

where T_{sw} is the switching cycle. Then, the input power can be derived from the input voltage multiplied by the input current where V_{ac} is the amplitude of the ac input voltage.

Equation (5) shows an important fact that the input power is inversely proportional to the operating frequency since $V_{ac}(t)$ is given and $i_{avg}(t)$ in (4) is inversely proportional to the operating frequency.

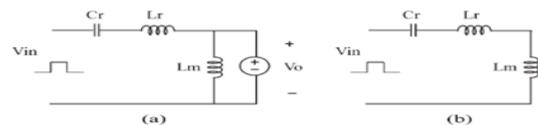


Fig.2.2.1 : Equivalent Circuit Of LLC Resonant Circuit
 (a).Secondary Sides Diodes Turn On (b).Secondary Sides Diodes Turn Off

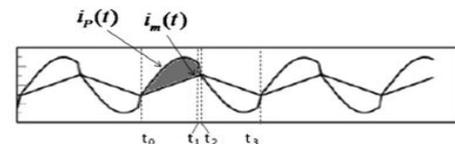


Fig.2.2.2 : waveform of ideal LLC converter ($f_{sw} > f_r$)

2.3. LLC CONVERTER STAGE

In this section, the goal is to find out the relationship between the operating frequency and the output power of the LLC converter. The previous works in [8]–[10] on a single-stage LLC converter did not derive the closed form of the LLC stage because of two major concerns.

First, the transient analysis is too complex. In a resonant circuit, the resonant current can be derived by the given driven source and loading condition. In the case of the LLC converter, the input voltage is a square waveform, and the output is a rectifier which behaves like a constant voltage load and behaves like an open circuit when two diodes are all reverse biased. The waveforms of the LLC converter are shown in Fig.2.1.2, where the operating frequency is larger than the resonant frequency in Fig.2.2.1 and the operating frequency is smaller than the resonant frequency. The primary side resonant current $i_p(t)$ can be substituted into the following equation to get the output power of the LLC converter:

$$P_{LLC} = \frac{E_{out}}{\frac{T_{sw}}{2}} = \frac{2 \int_{t_0}^{t_1} V_0(i_p(t)i_m(t)) dt}{T_{sw}} \dots (6)$$

where E_{out} is the output energy in a half period and $(i_p(t) - i_m(t))$ is the shadow area in Fig.2.2.2. However, the resonant current derivation is too complex which makes transient not suitable for system analysis.

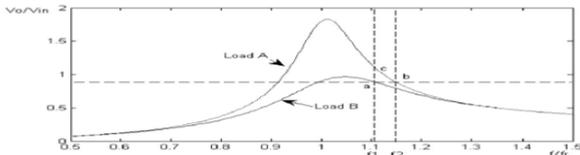


Fig .2.3.1 : Voltage Gain Versus Normalized Operating Frequency

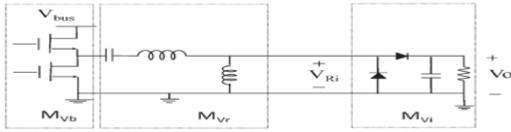


Fig .2.3.2 : Voltage Gain of the LLC Converter

The second consideration is that, in a dc/dc converter, the load is not constant, which may affect the system output power behaviors. However, since the system is frequency controlled, the frequency varies with loading condition so that there could be found load impedance corresponding to a given operating frequency. Fig. 9 explains the concept. In a dc/dc converter, the operating frequency varies with loading condition. When the load moves from load B to loadA, the operating frequency shifts from f_1 to f_2 correspondingly. That is, given a voltage gain and an operating frequency, corresponding load impedance can be found.

AC equivalent analysis helps to find out the corresponding load impedance. Fig .2.3.2 shows a typical ac circuit analysis. The system is divided into three parts, and the voltage gain is calculated separately.

The gain can be gotten as

$$M_v = \frac{V_0}{V_{bus}} = M_{vb} M_{vr} M_{vi} = \frac{V_i}{V_{bus}} \frac{V_{ri}}{V_i} \frac{V_0}{V_{ri}}$$

$$= \frac{\sqrt{2}}{\pi} \frac{1}{\sqrt{(1+A)^2 [1 - (\frac{\omega_0}{\omega})^2]^2 + \frac{1}{Q_L^2} (\frac{\omega}{\omega_0} \frac{A}{A+1} - \frac{\omega_0}{\omega})^2} b^2} \frac{\sqrt{2}}{\pi N} \quad (7)$$

where ω_o is the corner frequency $\frac{1}{\sqrt{L_r C_r}}$ A is the ratio of the leakage inductor to the magnetic inductor $\frac{L_r}{L_m}$, and Q_r is the loaded quality factor at the corner frequency

$$Q_L = \frac{\pi^2 n^2 R_L}{8} \sqrt{\frac{C_r}{(L_r + L_m)}} \quad \dots (8)$$

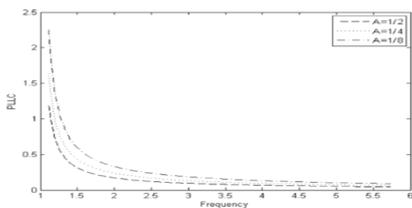


Fig.2.3.3 : PLLC versus frequency ($f_{sw} > f_r$)

Combining (7) and (8), the corresponding impedance RL can be derived.

Since the load impedance is gotten, the output power of the LLC converter can be derived easily as

$$P_{LLC} = \frac{V_0^2}{R_L} \quad \dots (9)$$

Substituting RL into (9), the output power of the LLC converter is

Fig.2.3.4 helps to realize the equation. The voltage gain is set to half of the gain at resonant frequency G_{fr} (the voltage gain at resonant frequency is constant under varied loading condition). The larger magnetic inductor L_m brings a flatter curve in the figure. The output power decreases when the operating frequency is lower than the resonant frequency. That is because the calculated corresponding impedance increases when the operating frequency decreases.

To have the same curve trend as the boost stage (output power increase with operating frequency) below resonant frequency, the voltage gain has to be set larger than G_{fr} . In Fig.2.3.2, the voltage gain is set to $1.1G_{fr}$. When the frequency is around resonant frequency, no load impedance corresponding to the operating frequency could be found because no loading condition can reach $1.1G_{fr}$ around resonant frequency. Therefore, the output power reduced to zero in Fig.2.3.5

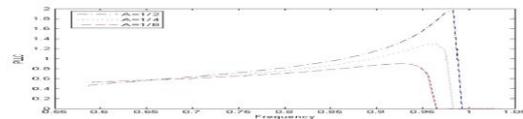


Fig.2.3.4 : PLLC versus frequency ($f_{sw} < f_r$)

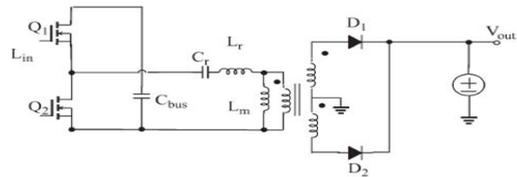


Fig.2.3.5: Simulation Model

In this paper, circuit simulation helps to find out the exact design parameters since there are errors between linear ac analysis and real circuit behaviors. Fig.2.3.5 shows the used simulation circuit. The load is replaced by a constant voltage source.

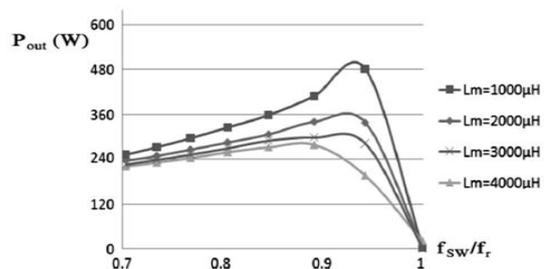


Fig.2.3.6: output power simulation ($f_{sw} < f_r$)

In Fig.2.3.6 are the simulation results which show the relationship between output power and frequency when the switching frequency f_{sw} is less than the resonant frequency f_r . In the period where the curve slope is negative, the input impedance of the LLC converter is inductive, and the ZVS can be achieved. The curve slope can be tuned by modifying the ratio of the magnetic inductor to the leakage inductor. In Fig. 9, C_r and L_r are given as $0.1 \mu F$ and $250 \mu H$, respectively, and the magnetic inductor varies from 1000 to 4000 μH .

Fig.2.3.7 shows the relationship between output power and frequency ($f_{sw} > f_r$). Larger L_m let the curve more flat. It helps to match the curve of the boost stage. The ratio of C_r to L_r is modified by doubling C_r and decreasing L_r by half in Fig. 15, which shows that the relationship between the output power and system switching frequency is the same except that P_{out} is doubled. That means that the power transferred into the system load doubles as the ratio of C_r to L_r doubles. Therefore, the system output power can be decided by modifying the ratio of C_r to L_r

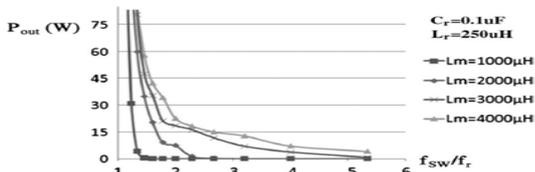


Fig.2.3.8: output power simulation ($f_{sw} > f_r$)

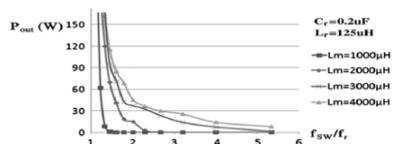


Fig.2.3.9: output power simulation after modifying the ratio of inductor to capacitor ($f_{sw} > f_r$)

III. LLC RESONANT CONVERTER

Three traditional resonant topologies were analyzed in above part. From the results, we can see that all of them will see big penalty for wide input range design. High circulating energy and high switching loss will occur at high input voltage. They are not suitable for front end DC/DC application. Although above analysis give us negative results, still we could learn something from it: For a resonant tank, working at its resonant frequency is the most efficient way. This rule applies to SRC and PRC very well. For SPRC, it has two resonant frequencies. Normally, working at its highest resonant frequency will be more efficient. To achieve zero voltage switching, the converter has to work on the negative slope of DC characteristic.

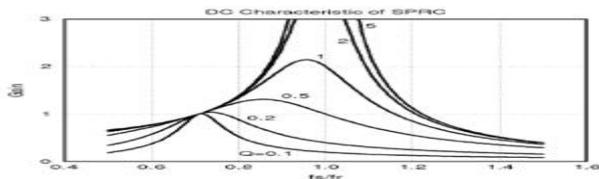
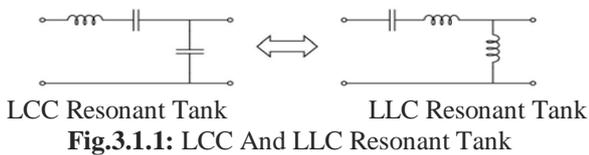


Fig.3.1.2: Dc Characteristics Of LCC Resonant Converter

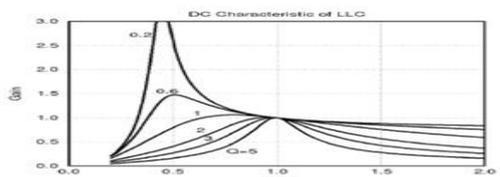


Fig.3.1.3: Dc Characteristics Of LLC Resonant Converter

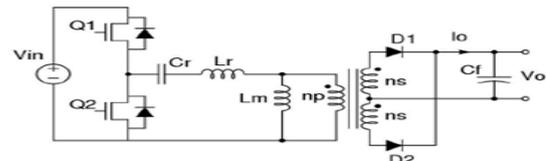


Fig.3.1.4: Half Bridge LLC Resonant Converter.

3.1 OPERATION OF LLC RESONANT CONVERTER

The DC characteristic of LLC resonant converter could be divided into ZVS region and ZCS region as shown in Fig.3.1.3. For this converter, there are two resonant frequencies. One is determined by the resonant components L_r and C_r . The other one is determined by L_m , C_r and load condition. As load getting heavier, the resonant frequency will shift to higher frequency. The two resonant frequencies are:

$$f_{r1} = \frac{1}{2\pi\sqrt{L_r C_r}} \quad \dots (10)$$

$$f_{r2} = \frac{1}{2\pi\sqrt{(L_m + L_r) C_r}} \quad \dots (11)$$

With this characteristic, for 400V operation, it could be placed at the resonant frequency of f_{r1} , which is a resonant frequency of series resonant tank of C_r and L_r . While input voltage drops, more gain can be achieved with lower switching frequency. With proper choose of resonant tank, the converter could operate within ZVS region for load and line variation.

From above discussion, the DC characteristic of LLC resonant converter could be also divided into three regions according to different mode of operation as shown in Figure 20. Our designed operating regions are region 1 and region 2. Region 3 is ZCS region. The converter should be prevented from entering region 3. In fact, there are many other operating modes for LLC resonant converter as load changes.

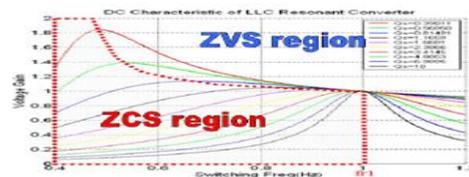


Fig.3.1.5: Dc Characteristics of LLC Resonant Converter

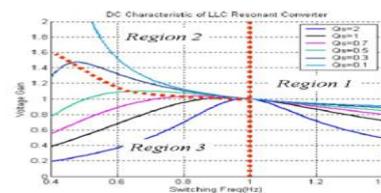


Fig.3.1.6: Three Operating Region of LLC Resonant Converter

In region 1, the converter works very similar to SRC. In this region, L_m never resonates with resonant capacitor C_r ; it is clamped by output voltage and acts as the load of the series resonant tank. With this passive load,

LLC resonant converter is able to operate at no load condition without the penalty of very high switching frequency. Also, with passive load L_m , ZVS could be ensured for any load condition. Here the operation will not be discussed in detail. There are several other modes of operation for light load condition.

Mode 1 (t_0 to t_1):

This mode begins when Q2 is turned off at t_0 . At this moment, resonant inductor L_r current is negative; it will flow through body diode of Q1, which creates a ZVS condition for Q1. Gate signal of Q1 should be applied during this mode. When resonant inductor L_r current flow through body diode of Q1, I_{Lr} begins to rise, this will force secondary diode D1 conduct and I_o begin to increase. Also, from this moment, transformer sees output voltage on the secondary side. L_m is charged with constant voltage.

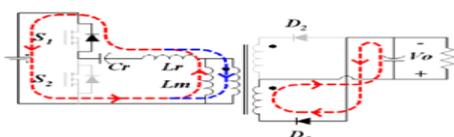


Fig.3.1.7: Circuit Diagram During Mode 1 In Region 2 Mode 2 (t_1 to t_2)

This mode begins when resonant inductor current I_{Lr} becomes positive. Since Q1 is turned on during mode 1, current will flow through MOSFET Q1. During this mode, output rectifier diode D1 conduct. The transformer voltage is clamped at V_o . L_m is linearly charged with output voltage, so it doesn't participate in the resonant during this period. In this mode, the circuit works like a SRC with resonant inductor L_r and resonant capacitor C_r . This mode ends when L_r current is the same as L_m current. Output current reach zero.

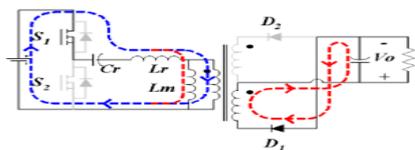


Fig.3.1.8: Circuit Diagram during Mode 2 In Region 2 Mode 3 (t_2 to t_3)

At t_2 , the two inductor's currents are equal. Output current reach zero. Both output rectifier diodes D1 and D2 is reverse biased. Transformer secondary voltage is lower than output voltage. Output is separated from transformer. During this period, since output is separated from primary, L_m is freed to participate resonant. It will form a resonant tank of L_m in series with L_r resonant with C_r . This mode ends when Q1 is turned off. As can be seen from the waveform, Q1 turn off current at t_3 is small compare with peak current. For next half cycle, the operation is same as analyzed above

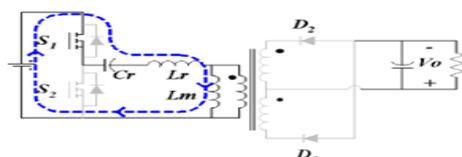


Fig.3.1.9: Circuit Diagram During Mode 2 In Region 2

IV. DESIGN PROCEDURE

Based on the analysis shown before, a design procedure is presented in the following. A 75-W 110-Vac-to-18-Vdc dc/dc converter is taken as a design example.

- 1) First, set the target voltage of the bus capacitor. Chosen in this case for DCM is 390 V.
- 2) Decide the minimum operating frequency. Higher operating frequency brings higher switching loss. Therefore, 50 kHz is chosen as the minimum operating frequency.
- 3) Given the bus voltage and the full-load switching frequency, input inductor L_{in} can be gotten by (4) and (5) as $560 \mu\text{H}$.
- 4) Select the resonant frequency of the LLC converter and the ratio of the magnetic inductor to the leakage inductor. To match P_{boost} and $PLLC$, the smooth region of the curves in Figs. 11 and 14 is used. The curves are flatter when the ratio of the leakage inductor to the magnetic inductor is lower. On the other hand, the curves are much flat at the high-frequency region. However, being close to resonant frequency means higher efficiency. Finally, $1.2f_r$ is chosen as the minimum operating frequency of the LLC converter, and the ratio of the magnetic inductor to the leakage inductor is set to eight. Therefore, the resonant frequency can be calculated as 41.6 kHz.
- 5) Decide the LLC series capacitor. The relationship between system power capacity and series capacitor is shown in Figs. 14 and 15. Then, the series capacitor is chosen as $0.1 \mu\text{F}$. Then, the leakage inductor can be calculated as $175.2 \mu\text{H}$ by the equation
- 6) To avoid operating at extremely high frequency, the system enters burst mode while the operating frequency reaches triple the minimum operating frequency.

TABLE 4.1. Designed Parameters.

Input voltage	110V _{ac}
Output voltage	18V _{dc}
Output power	75 W
V _{cbus}	320V
Minimum switching frequency	50kHz

TABLE 4.2. Component Parameters

L_r	135 μH	L_{in}	180 μH
L_m	1.13mH	NF_p	45 turns
C_r	0.1 μF	N_s	4 turns
C_{bus}	180 μF *2	Power mos	STP25NM50N
C_{out}	470 μF *2	Rectify diode	STP30S45CT

V. SIMULATION RESULTS

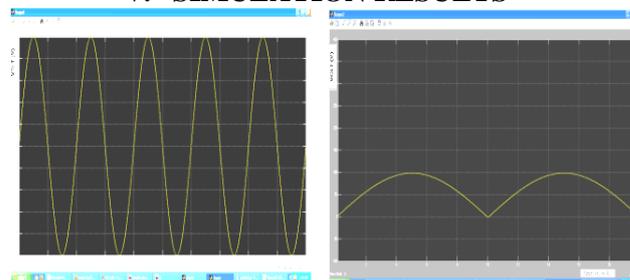


Fig.5.1: Ac input voltage Fig.5.2: Rectified half wave dc voltage

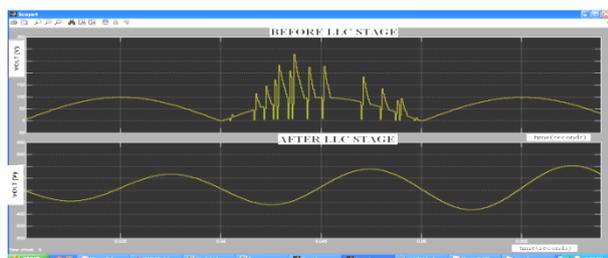


Fig.5.3: LLC output voltage



Fig.5.4: Voltage Across Switches

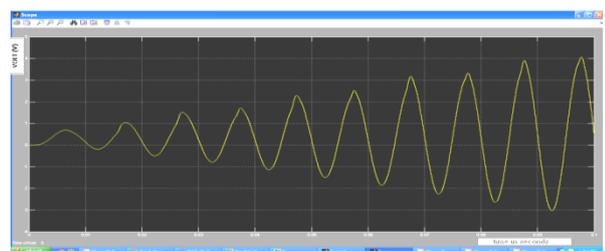


Fig.5.5: Load Voltage

VI. CONCLUSION

Analysis of a single-stage LLC converter has been presented in this paper. To stabilize the bus capacitor voltage, the power balance of the boost stage and the LLC converter has been investigated, and then, the design procedure has been presented. For verification, a 75-W 110-Vac-to-18-Vdc single-stage converter has been designed and implemented. The bus voltage can be kept around 390 V with a maximum variation of 38 V. The system efficiency can reach 90% while the loading condition is heavier than 20%.

REFERENCES

- [1] G.Moschopoulos and P. Jain, "Single-phase single-stage power-factor corrected converter topologies," *IEEE Trans. Ind. Electron.*, vol. 52, no. 1, pp. 23–35, Feb. 2005.
- [2] Q. Mei, G. Moschopoulos, H. Pinheiro, and P. Jain, "Analysis and design of a single stage power factor corrected full-bridge converter," in *Proc. Appl. Power Electron. Conf. Expo.*, Mar. 1999, pp. 119–125.
- [3] F. Zhang and C. Gong, "A new control strategy of single-stage flyback inverter," *IEEE Trans. Ind. Electron.*, vol. 56, no. 8, pp. 3169–3173, Aug. 2009.
- [4] Y.-M. Liu and L.-K. Chang, "Single-stage soft-switching AC–DC converter with input-current shaping for universal line applications," *IEEE Trans. Ind. Electron.*, vol. 56, no. 2, pp. 467–479, Feb. 2009.
- [5] S.Moo, K. H. Lee, H. L. Cheng, and W.M. Chen, "A single-stage highpower-factor electronic ballast with ZVS buck-boost conversion," *IEEE Trans. Ind. Electron.*, vol. 56, no. 4, pp. 1136–1146, Apr. 2009.
- [6] L. Yan and G. Moschopoulos, "A single-stage AC–DC forward converter with input power factor correction and

reduced DC bus voltage," in *Proc. Telecommun. Energy Conf.*, Oct. 2003, pp. 132–139.

- [7] P. Das, S. Li, and G. Moschopoulos, "An improved AC–DC singlestage full-bridge converter with reduced DC bus volt," *IEEE Trans. Ind. Electron.*, vol. 56, no. 12, pp. 4882–4893, Dec. 2009.
- [8] M. S. Agamy and P. K. Jain, "A three-level resonant single-stage power factor correction converter: Analysis, design, and implementation," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2095–2107, Jun. 2009.
- [9] D. D.-C. Lu, H. H.-C. Lu, and V. Pjevalica, "Single-stage AC/DC boost– forward converter with high power factor and regulated bus and output voltages," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2128–2132, Jun. 2009.
- [10] C. M. Lai, R. C. Lee, T.W.Wang, and K. K. Shyu, "Design and implementation of a single-stage LLC resonant converter with high power factor," in *Proc. IEEE Int. Symp. Ind. Electron.*, Jun. 2007, pp. 455–460.
- [11] B. C. Kim, K. B. Park, C. E. Kim, B. H. Lee, and G. W. Moon, "LLC resonant converter with adaptive link-voltage variation for a high-powerdensity adapter," *IEEE Trans. Ind. Electron.*, vol. 25, no. 9, pp. 2248–2252, Sep. 2010.
- [12] STMicroelectronics, Application Note LLC Resonant Half-Bridge Converter Design Guideline, Mar. 2007.
- [13] F. Dianbo, K. Pengju, W. Shuo, F. C. Lee, and X. Ming, "Analysis and suppression of conducted EMI emissions for front-end LLC resonant DC/DC converters," in *Proc. Power Electron. Spec. Conf.*, Jun. 2008, pp. 1144–1150.



1. P.Ganesh was born in Tamil Nadu, India, on November 24, 1988. He was received the B.E. Degree in Electrical and Electronics engineering Branch from Sri

Subramanya College of Engineering and Technology, Palani, Tamilnadu in 2010. He is currently pursuing the M.E Degree in Power Electronics and Drives at Sri Subramanya College of Engineering and Technology, Palani, Affiliated to Anna University Chennai, India. His research topics include power electronics application to power system, power quality, special electrical machines and power electronics.



2. T. Manokaran was born in Tamil Nadu, India, on July 19 1974. He received the A.M.I.E. and M.E. degrees in electrical engineering Branch from Government

College of Technology, Coimbatore, India. in 2006 currently, he is pursuing the Ph.D. degree at Anna University of Technology Madurai, Madurai, India. His research topics include power electronics application to power system, power quality, special electrical machines and power electronics. He is currently working as associate professor in Electrical and Electronics Engineering Department at Sri Subramanya College of Engg Technology, Palani Dindigul (Dt) Affiliated to Anna University, Chennai, Tamilnadu, India.

An Enhancement of Peak to Average Power Ratio Reduction in OFDM Using CAP-PT Method

C. Raja Rajeshwari¹, K. Manojkumar²

Assistant Professors Department of ECE

¹ Karur Engineering College, Karur

² Christ The King Engineering College, Karamadai Tamil nadu, INDIA

Abstract: The main drawback of OFDM signals demonstrates high fluctuations termed as Peak to Average Power Ratio (PAPR). A signal with high PAPR when amplified by a Power Amplifier (PA) without any special treatment results in severe impairments, i.e., non-linear distortions. Hence this non linearity destroys the orthogonality of the OFDM signal and introduces out-of-band radiation and in band distortions. The existing method doesn't have high capability of PAPR reduction with significant high Bit Error Rate (BER) degradation. Hence this proposed CAP-PT (clipping, Amplification and Peak Windowing/ Partial Transmit) method combines the effect of clipping the peak signals and amplifies the signals that have low amplitude. Out of band radiations which is caused by clipping is suppressed by peak windowing process. Then followed by partial transmit which reduces in band distortions. The PAPR reduces to the great extent without BER degradation. Simulation results demonstrate that the proposed CAP-PT method yields the better performance than the existing methods.

Key-words: OFDM, PAPR Reduction, Multicarrier transmission, Clipping, Amplification, Peak Windowing, Partial Transmit

I. Introduction

OFDM is one of the multicarrier modulation (MCM) techniques for 4th Generation (4G) wireless communication. This technique is very attractive technique for high-speed data transmission used in mobile communication, Digital terrestrial mobile communication, Digital Audio Broadcasting (DAB), Digital Video Broadcasting terrestrial (DVB-T), wireless asynchronous transfer mode (WATM), Modem/ADSL. OFDM has many advantages such as robustness in frequency selective fading channels, high spectral efficiency, immunity to inter-symbol interference and capability of handling very strong multipath fading [1-3].

But OFDM is having major drawback of a high Peak-to-Average Power ratio (PAPR) [4-5]. This causes clipping of the OFDM signal by the High power amplifier (HPA) and in the HPA output producing nonlinearity. This non-linearity distortion will result in-band distortion and out-of-band radiation. The in-band distortion causes system performance degradation and the out-of-band radiation causes adjacent channel interference (ACI) that affects systems working in neighbor band. Hence the OFDM signal may have In-band and Out-of-band distortion which degradation of Bit-error-rate (BER) performance. One solution is to use a linear power amplifier with large dynamic range. However, it has poor efficiency and is expensive too.

This paper continues as follows: Section 2 explains the concept of OFDM. Various approaches of PAPR reduction is described in Section 3. Section 4 explains the PAPR. Proposed method is explained in section 5. Simulations Result was discussed in section 6. Conclusion and future work explained in section 7.

II. Orthogonal Frequency Division Multiplexing (OFDM)

OFDM is a Multicarrier Transmission technique which divides the available spectrum into many carriers each one being modulated by a low data rate stream. OFDM is similar to Frequency Division Multiple Access (FDMA) in that the multiple user access is achieved by sub-dividing the available bandwidth into multiple channels, which are then allocated to users. This is achieved by making all the carriers orthogonal to one another, preventing interference between them [6, 7].

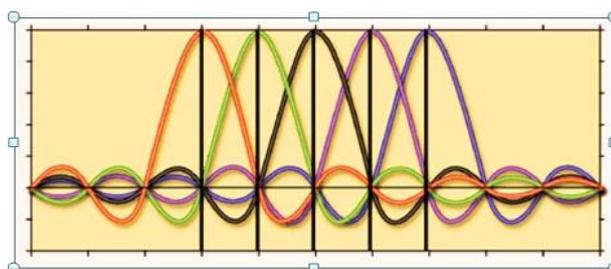


Fig.1 Orthogonality of subcarriers

The subcarriers in an OFDM signal are spaced close as is theoretically possible which maintain orthogonality between them. The orthogonality of the carriers means that each carrier has an integer number of cycles over a symbol period. Due to this the spectrum of each carrier has a null at the center frequency of each of the other carriers in the system. This results in no interference between the carriers, allowing them to be spaced as close as theoretically possible.

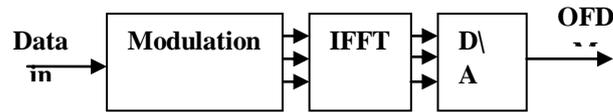


Fig.2 OFDM transmitter

To generate OFDM successfully the relationship between all the carriers must be carefully controlled to maintain the orthogonality of the carriers. For this reason, OFDM is generated by firstly choosing the spectrum required, based on the input data, and modulation scheme used. Each carrier to be produced is assigned some data to transmit. Fig.2 shows the transmission of OFDM signal. The orthogonal carriers required for the OFDM signal can be easily generated by setting the amplitude and phase of each frequency bin, then performing the IFFT. Since each bin of an IFFT corresponds to the amplitude and phase of a set of orthogonal sinusoids, the reverse process guarantees that the carriers generated are orthogonal.

III. Related Works

Several algorithms [8-11] have been proposed to handle this PAPR problem. However, none of these algorithms have produced significant reduction of PAPR in OFDM systems. A simple Encodable /Decodable OFDM QPSK proposed in [12] used Reed-Muller code with QPSK. This could reduce the PAPR but it could not be used with higher order signal constellations. OFDM PAPR reduction by a rotation of redundancy bit position in subblock code word scheme was proposed in [13]. In this method the redundant bit positions of subblock code words are rotated and the lowest PAPR codeword is chosen by a feedback scheme. However, the side information for bit position is required. Companding transform [14,15] compresses as large signal while enhancing a small signal that can achieve a desired PAPR but with a significant increase in the bit error rate (BER). Selective Mapping (SLM) [16, 17] is based on the creation of P different signals from the original input data vector. The tone reservation method [18], also called Peak-Reduction Carriers reserves a set of subcarriers to create the PAPR reduction signal $c(t)$, while the tone injection method makes use of so-called expanded constellation diagram [19]. Block coding was proposed in [20], in which a data sequence is embedded in a longer sequence, and only a subset of all these possible sequences is used to exclude patterns generating high PAPR.

IV. The PAPR of OFDM System

An important limitation of OFDM is that it suffers from a high Peak-to-Average Power Ratio (PAPR) resulting from the coherent sum of several carriers. This forces the power amplifier to have a large input backoff and operate inefficiently in its linear region to avoid intermodulation products. High PAPR also affects D/A converters negatively and may lower the range of transmission. The PAPR of OFDM is defined as the ratio between the maximum power and the average power. The PAPR of the OFDM signal $X(t)$ is defined as

$$PAPR = \frac{P_{peak}}{P_{average}} = \frac{\max_n [|x_n|^2]}{E [|x_n|^2]} \quad (1)$$

Where x_n = An OFDM signal after IFFT (Inverse Fast Fourier transform)
 $E[.]$ = Expectation operator, it is an average power.

V. Proposed Method

This method is simple but efficient algorithm for PAPR reduction by using CAP-PT (Clipping, Amplification and Peak Windowing / Partial Transmit) method. Fig.3 depicts the block diagram of proposed CAP-PT model.

The algorithm is based on clipping the amplitude of OFDM signal that exceeds threshold value say (A). The amplitude of signals exceeding clipping threshold value (say A) is clipped and the signals having amplitudes less than amplification threshold value say (say B) is amplified to threshold level. PAPR is calculated for clipped signal and compared with that of the original PAPR of OFDM signal. Clipping can be performed by clipping the complex envelope of OFDM signal.

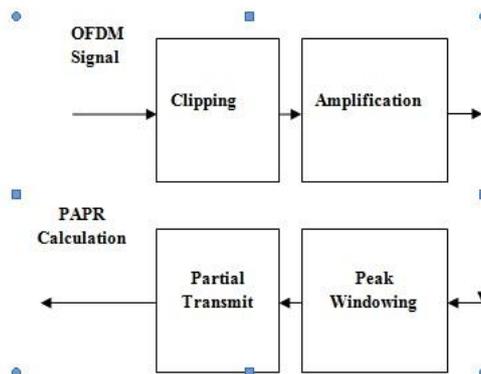


Fig.3 Block Diagram of Proposed CAP-PT model

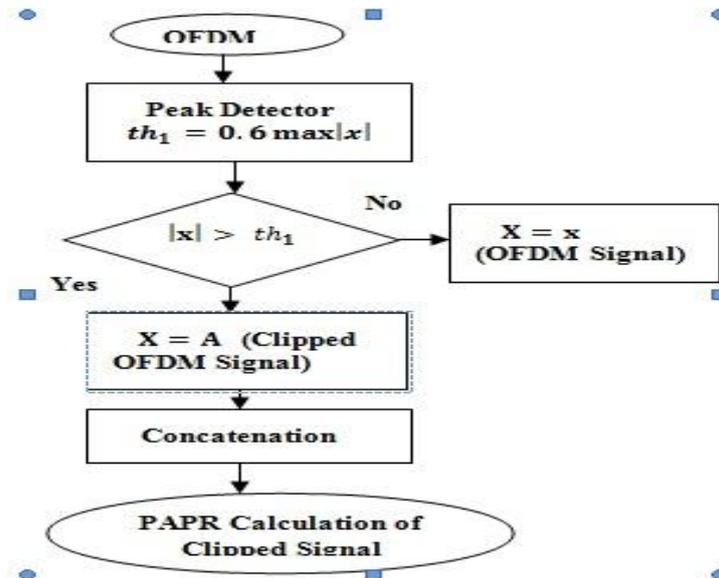


Fig.4 Flowchart for Clipping

Probability of clipping is defined as

$$CCDF = P_c = P \{A < X < \infty\} = e^{-\frac{A}{2\sigma^2}} \quad (2)$$

where, A is clipping level and σ^2 is variance of x

Taking ln on both sides we get,

$$-2 \ln(CCDF) = PAPR \quad (3)$$

Or

$$A^2 = -\ln(CCDF)\sigma^2 \quad (4)$$

Above equations gives the relation between CCDF and PAPR. This concludes that with decrease in PAPR the CCDF increases and vice versa.

This clipping and amplification process then combines with peak windowing technique. Fig.4 shows the flowchart for clipping process. The process of peak-windowing is realized by multiplying window function from the biggest peak to the smaller peaks in the peaks exceeding fixed value level. In proposed method, peak-windowing is performed on clipped and amplified OFDM signal. Our method suppresses the out of radiation caused by clipping which results in PAPR reduction. This technique does not require a high power amplifier which increases the circuit complexity at the receiver.

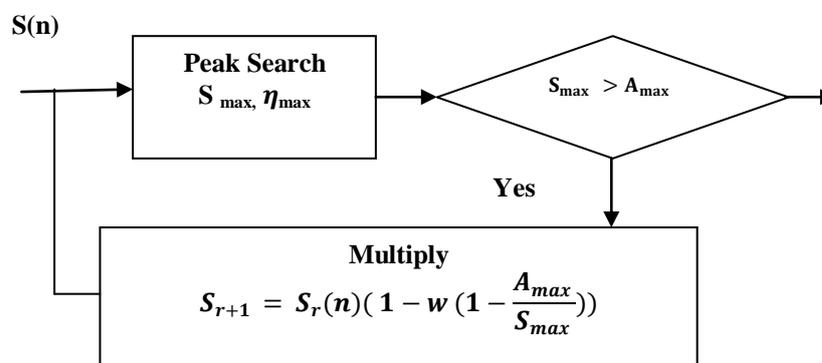


Fig.5. Process of Peak windowing

Fig.5 shows the process of peak windowing. s_{max} is the symbol which has the biggest peak in an iterative signal $x(n)$. n_{max} is its index. r is the iteration number. w is the window function. A_{max} is the maximum amplitude after clipping and Amplification. P_{mean} is the average input power of the OFDM signal before clipping.

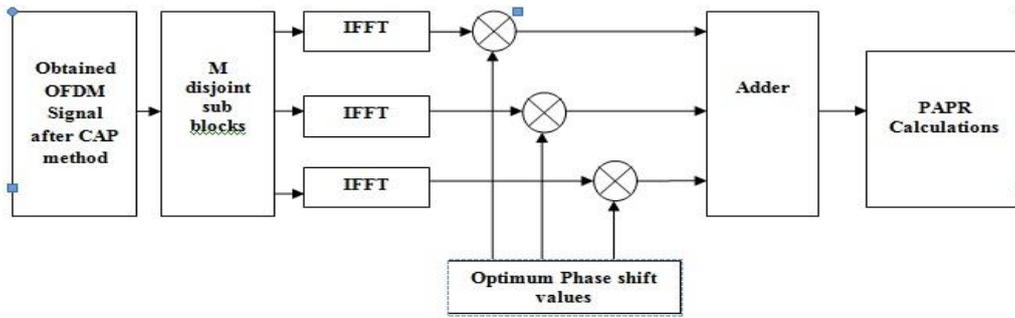


Fig.6. Block diagram of Partial Transmit

In Partial Transmit Sequences (PTS) method, the IFFT inputs symbols are divided into several frequency disjoint subblocks. The IFFT operation is performed not on the N subcarriers in total, but separately on these subblocks. The output of l -th subblock is then multiplied by so-called complex rotation factor b_l . The values of b_l for all l are then optimized to find complex rotation factors resulting in the lowest symbol PAPR. This optimization has to be performed in the real time for each IFFT input symbol. The information about used complex rotation factors has to be send as a side information to the receiver. This reduces the useful data rate. The Block diagram of PT is shown in the fig.6.

VI. Simulation Results and Discussions

Simulations of clipping and amplification methods and proposed CAP/ PT method are compared. The simulation parameters are listed in Table 1. OFDM signal is generated using QPSK modulation. Then it undergoes clipping process. In clipping, peak of the OFDM signal is detected by using peak detector (th_1) to set the clipping threshold (A). Then the peak of the OFDM which is above th_1 i.e., 0.6 times of peak value of the OFDM signal is clipped. Figure.7 shows the comparison of original signal and clipped signal. Threshold (th_1) is calculated by using peak detector as 0.45. After clipping process, amplification takes place which amplifies the signal that has low amplitude.

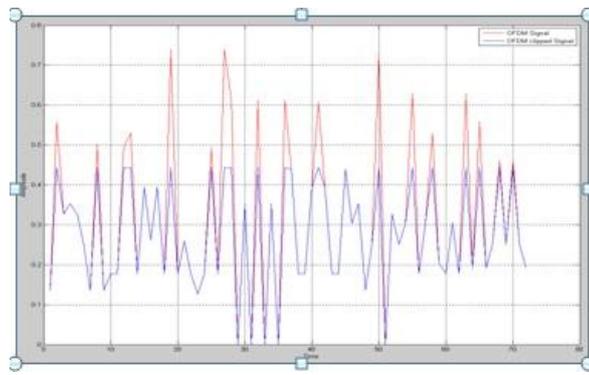


Fig.7 comparison of original and clipped OFDM signal

TABLE 1 simulation parameters

Modulation	M-QAM, M-PSK
Number of data subcarriers	64
Number of FFT points	8
Number of sub-blocks	4
Window function	Blackman
Channel Model	AWGN

In Peak detection, components of the amplitude which exceeds fixed value is detected. Peak-detection is transferred to frequency domain by FFT. Out of band components which are generated by FFT are set to zero. Thus out of band radiations is suppressed by peak windowing. As PT is distortion less technique, it reduces in band distortion. PAPR reduced to 2.8 dB when CAP combined with PT method.

Comparison between different M-ary (M- QAM, and M-PSK, M=16, 32, 64) modulation technique for CAP-PT method is investigated choosing clipping threshold to be 60 % of the peak volt of OFDM signal. Fig. 8 shows the comparison between 16-PSK and 16- QAM, 32-PSK and 32- QAM, 64-PSK and 64-QAM respectively. On X-axis different modulation techniques (M-QAM, and M-PSK, M=16, 32, 64 as 1, 2, 3, 4, 5 & 6 respectively) and on Y-axis PAPR reduction

(in dB) is plotted. It is observed that with increase in M (M=16, 32, 64) the PAPR increases and so does PAPR reduction. It is also observed that PSK has better performance over QAM in terms of PAPR reduction i.e. PSKs have more PAPR reduction than QAMs.

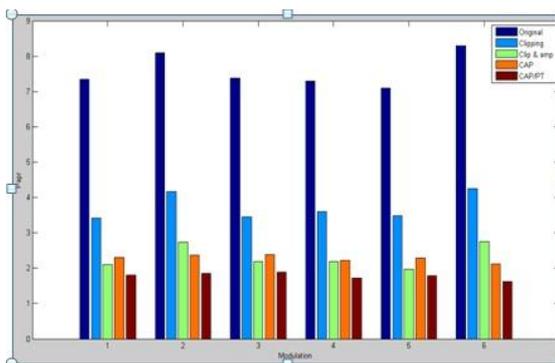


Fig. 8 Comparison between 16- QAM, 32- QAM, 64-QAM and 16 –PSK, 32-PSK, 64-PSK respectively

The value of PAPR reduction (in dB) is indicated in the figure. The conclusion drawn from this is that the PSK modulation technique is better than QAM in OFDM in terms of PAPR and QAM. To say in other way, if PSK modulation is used during generation of OFDM signal, PAPR will be less and if PAPR occurred is high(since the PAPR is dependent of the bit sequence), PAPR can be considerably decreased using CAP-PT method.

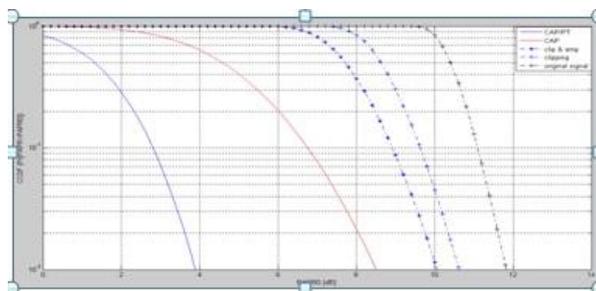


Fig.9. comparison of PAPR with probability of clipping of CAP-PT method

Figure.9 shows the plot of PAPR Vs complementary cumulative distributive function (CCDF) of clipping, amplification and peak windowing combined with Partial transmit methods. Table 2 shows that the increase in the clipping probability, PAPR decreases.

Table 2 Reduction of PAPR of various methods

Method	Original Signal	Clipped Signal	Clipped and Amplified Signal	CAP	CAP/PT
CCDF					
0.01	11.8	10.8	10	8.5	3.9
0.1	11	9.8	9	7	2.8

VII. Conclusion And Future work

In this paper Orthogonal Frequency Division Multiplexing (OFDM) is presented as one of best candidates for fourth generation (4G) communication system, as it can combat radio impairments very effectively with very efficient use of spectrum. But every advantage comes with some disadvantages. Different disadvantages of OFDM such as frequency synchronization, maintaining orthogonality among carriers, local oscillator offset etc.

This paper mainly focuses on one of the challenging drawback of OFDM that is high peak to average power ratio (PAPR). A new CAP-PT method is proposed in this paper. Clipping is a non-linear process and causes significant inband distortion that causes bit error degradation, outband distortion causes degradation in spectral efficiency. The drawback of clipping can be combated using peak windowing technique. Peak windowing seems to be beneficial. Partial transmit significantly reduce in band distortions with large PAPR reduction.

References:

- [1] DVB, "Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)," ETSI EN 302755 V1.1.1 September 2009.
- [2] Kyeongcheol Yang and Seok-II Chang, "Peak-To-Average Power control in OFDM using standard Arrays of Linear Block Code, IEEE communication letters, Vol.7, No.4, pp. 174-176, April 2003
- [3] R. Van Nee and R. Prasad, OFDM for Wireless Multimedia Communications, Artech House, Boston, Mass, USA, 2000.
- [4] A.Ghassemi and T. A. Gulliver, "Fractional Selective mapping using Decimation in time IFFT/FFT.", Publication in WCNC 2008 Proceedings IEEE, pp. 543- 547.
- [5] Abdulla A. Adouda, "PAPR Reduction of OFDM signal using Turbo coding and Selective Mapping.", Proceedings of 6th Nordic signal processing symposium-NORSIG2004.
- [6] Hujun Yin and Siavash Alamouti. "OFDMA: A Broadband Wireless Access Technology". IEEE Sarnoff Symposium, August 2007.
- [7] Paiement.R.V, "Evaluation of Single Carrier and Multicarrier Modulation Techniques for Digital ATV Terrestrial Broadcasting CRC Rep", Ottawa, ON, Canada, CRC-RP-004, 2004.
- [8] Mario Bonaccorso, Veronic Buzenac. "Reducing the Peak to Average Power Ratio in OFDM Systems", Dix-septième colloque GRETSI, Vannes, september 2008.
- [9] Han.S.H, and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," IEEE Wirelless Communications Magazine, Vol.12, No.2, pp.56-65, 2005.
- [10] Telladoj, "Peak to average power reduction for multicarrier modulation", Ph.D. distortion, standard University, 2000.
- [11] Rana.M.M "Clipping based PAPR reduction method for LTTE OFDMA Systems", IEEE IJECS-IJENS, Vol: 10 No.:05, 2000.
- [12] ChongC., V. and T arokh,V, "A Simple Encodable / Decodable OFDM QPSK Code with Low Peak-to-Mean Envelope Power Ratio. IEEE Transactions on Information Theory, Yol. 41, No 7, pp. 3025-3029, 2001.
- [13] Tan, M., et al., "OFDM Peak Power Reduction by a Novel Coding Scheme with Threshold Control, IEEE Vehicular Technology Conference Y, ol. 2, No 54ND, pp. 669-612,2001.
- [14] Wang X, Tjhung, T.T., Ng, C.S., "Reduction of peak to average power ratio of OFDM system using a companding technique", IEEE transactions on broadcasting, Vol.25, Sep 2002.
- [15] Huang,X ., et al., "Companding Transform for the Reduction of Peak-to-Average Power Ratio of OFDM Signals, IEEE Vehicular Technology Conference Vol. 2, No 53, pp. 835-839,2001.
- [16] Bauml, R.W., Fischer, R.F.H., and Huber, J.B." Reducing the peak to average power ratio of multi carrier modulation by selective mapping", IEEE Electronics Letters, Vol.32, Oct 2006.
- [17] Han.S.H and J.H. Lee, "Modified selected mapping technique for PAPR reduction of coded OFDM signal," IEEE Transactions on Broadcasting, Vol.50, No.3, pp.335-341,2004.
- [18] I. M. Mahafeno, Y. Lou'et, and J. F. H'elard, "Peak-to-average power ratio reduction using second order cone programming based tone reservation for terrestrial digital video broadcasting systems," IET Communications, vol. 3, no. 7, pp. 1250-1261, 2009.
- [19] TELLADO, J. "Multicarrier Modulation with Low PAR, Kluwer Academic Publishers, 2000
- [20] Wilkison, T.A. and Jones, A.E., "Minimization of the peak to mean envelope power ratio of multicarrier transmission schemes by block coding", IEEE Vehicular Technology Conference, Vol.2., Jul 2005.

Bibliography:



Raja Rajeshwari.C received the B.E. from Kongu Engineering College, Perundurai affiliated to Anna University, Chennai in 2009. She is currently working towards the M.E. degree at Sri Shakthi Institute of Engineering And Technology, Affiliated to Anna University of Technology, Coimbatore. She has 14 months of teaching experience. Her current research interests include the areas of wireless communications, especially for OFDM systems with emphasis on research of PAPR reduction.



M. Raj Kumari was born in Tamil Nadu, India in the year 1973. She has completed B.E., in Electronic and Communication Engineering in the year 1995 at Thiagarajan College of Engineering, Tamil Nadu, India and M.E., in Communication Systems in the year 1997 from the same Institution. She is working as Lectutrer since 1998 till date. She has an overall teaching experience of 14 years in various Engineering institutions both in India and Malaysia. At present she is working as Senior Assistant professor in ECE dept in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Her Areas of Interest are Wireless Communication, Microwave Engineering and Terahertz Communication. She is a member of IEEE.

Application of Soft Computing Techniques for Factory Floor Automation to Digital Ecosystems

P VL Narayana Rao

Department of Computer Science and Engineering,
Institute of Technology,
Ambo University, Ambo, Ethiopia

Abstract: This Research paper presents from the Process Definition Tools implementation, which is an automated manufacturing testbed that has been integrated into a Service Orientated Architecture (SOA). The enablement of this SOA in terms of Digital Business Ecosystem (DBE) research presents new applications to Small to Medium Enterprises (SMEs) that support, require and supply automated manufacturing systems. These new applications and characteristics in the context of Digital Ecosystem tools, technologies and general architectural changes are examined. In particular new opportunities in virtualization of automation and integration at both lower and higher levels of distributed applications are presented in a transformed view of a DBE.

Key Words: Automated Manufacturing, SOA, DBE, SME, Ecosystem

I. INTRODUCTION

Web Services and hardware development have combined to place factory floor level devices within the reach of large scale distributed computing applications. In terms of new SOA development this new dimension in potential services presents new possibilities in Digital Ecosystems. These can be seen to stem from increased application pervasiveness via the SOA enablement of the Factory floor that will introduce new services and service providers into DBE based applications.

Investigating this new DBE enablement this paper examines the new structures that could be formed in DBE's terms of the new users, service providers and applications. In particular these new DBE features are proposed and placed within the context of typical DBE architectures. This is done using experience from research projects into SOA enablement of factory floor level devices. The paper uses a three layer examination of the impact of factory floor automation on the DBE. Initially the advances in enabling devices from the perspective of the manufacturer are presented. This examines the technology that is to enable SOA based automation and how it is likely to be deployed. The new functions that this enablement presents to the DBE is examined, with a focus on new application areas likely to effect the future structures of DBE's using factory floor level services.

II. RELATED WORK

A.. Motivation behind the Web Services (WS) enablement of PLC's (Programmable Logic Controllers) and factory floor elements.

Manufacturers are seeking more agile, flexible and adaptive approaches to their production systems as well as to encompassing customers as a part of the business process

[1]. The modern production system model must be able to cope with these changes and comply with small quantities of highly customised design-to-order products, where additional services and value-added benefits like product upgrades and future reconfigurations are as important as the product itself. [2].

The need for a SOA enablement of devices on the factory floor has emerged from research into the improvement of automation of factory process. In particular within advanced western economies the need to have flexible production lines able to specify and react effectively to change is a key factor in enabling the production of high quality value added products that this sector is focused on. Within this research elements such as basic Robotics and manufacturing assembly has been the focus of work looking to introduce greater automation and reconfiguration [20].

The SOCRADES [3], SODA [4] and SIRENA [5] project within the EU have taken significant steps in the initial investigation of SOA enablement and support of factory floor level devices in distributed automation applications.

These projects are focused at the enablement of SOA within factory floor elements. As this enablement is starting to be defined and applied in test environments in these and similar projects the impacts on SOA based applications in general can begin to be analysed.

B. Significance to the Digital Business Ecosystem

There are more than 228000 manufacturing Small to Medium Enterprises (SMEs) in the EU, at the Lisbon summit in 2003 research efforts were agreed at focusing knowledge based manufacturing [6]. An important part of this was

the commitment of funds to support investigation into Digital Business Ecosystems in the EU. A digital business Ecosystem can commonly be seen as mechanism by which SMEs, are provided with applications consisting of distributed services normally out of their reach due to lack of resources or skills.

A popular application area has been within ICT based applications and knowledge sharing such to support supply chain management and specialised processing of data [7,8]. The key technical enabler of these distributed environments has been advances in web services and the provision of SOA based resource sharing. To date many DBE exist at what can be seen as the middleware of the distributed computing model, and often link applications and data.

Motivations behind the growth and evolution of DBE can be seen as centered in two main areas, these can be seen as technical and organisational. In terms of the

organization an main proponent of DBE development particularly in the EU has been governmental organisations. These DBE's are formed with the aim to promote the development of and sharing of skills in specific communities SME's. For example the EU's vision of DBE forms along socio economic boundaries to promote innovation and development in industries and localities improving European competitiveness [9,10]. This vision is not specifically technical but focused on supporting innovation in the distributed collaboration enabled by the DBE.

In order for DBE's to develop technically the new business models that they promote have attracted software development in both the commercial and open source community. This is often linked to genuine commercial motivation where the provision of software to support the DBE can present vendors like SAP with larger sets of users, developers and potential customers to DBE based products. An approach

that has presented companies like SAP with new business models, where an ecosystem could be built as SAP components around SAP products such as Net Weaver [11]. With the ecosystem enabling thousands of independent developers to start writing specialized programs that plug into the Net Weaver framework. The same model can also be seen to apply in the Open Source DBE vision. Where the enablement of a DBE using Open Source software can introduce new developers and users to the software [12].

Therefore the DBE like the development of SOA in factory automation is motivated by social economic goal of increasing skills and competitiveness in regional economies. Furthermore the technical enablement of the DBE like in factory automation promises new business opportunities for software providers. Combined within the DBE this combination of business and technical innovation will have a significant impact in DBE use and development.

To date within the field of industrial automation the enablement of factory floor elements the use of DBE type applications is yet to become an established commercial model. However as distributed applications emerge around the next generation of SOA based automation the DBE's is likely to adapt and form around this technology. Integration of this factory floor level data in many DBE's will give the applications greater depth and completeness.

III. ENABLING DEVICES

A. Device Profile for Web Services.

Nowadays globalisation, emergent technologies and customer demands require flexible operation on a global scale as well as agile, flexible and adaptive manufacturing control. In response to these rapid, continuous and unpredictable changes organisations should embrace a flexible, adaptive, collaborative, and responsive paradigm.

The key enabler is the integration of automation systems with other business entities including marketing, engineering, product design, business process management and personnel. However, the integration of a business enterprise with its shop floor and with business partners correctly is largely achieved via the use of proprietary complex systems due to disparate technologies being used in manufacturing systems and business enterprise [13]. In addition, lack of standardized technologies and interoperability are main obstacles in effectively integrating

the enterprise among business partners, thus hindering integration with DBEs.

Recently, the SOA and WS approach has been proposed by a number of researchers as mentioned earlier to enhance networked organisations to support the required dynamic networking of the inter- and intra- business collaborators into the factory floor environment.

This would enable this level of knowledge and information to be seamlessly shared and accessed through the use of standardised web technologies. Thus the use of SOA holds the potential to break the integration chains of proprietary and legacy software that is surrounding factory automation systems, and preventing them from joining effective DBEs.

The implementation technology to support the dynamic changes from the business and management level of integrating the shop floor is generally recognised will be achieved via SOA enabled by DPWS (Device Profile For Web Services) [14,15]. DPWS is a light-weight protocol stack that enables the service discovery and message passing to the higher local and remote control level.

The protocol can be implemented at control devices (PLC, Embedded devices). The protocol is defined with WS-, SOAP-XML, WSDL, UDDI, and TCP/IP/UDP/HTTP stacks and it is becoming the standard for the industrial network application and has its main implementations in the home electronics /consumer computing area [23]. This DPWS protocol will allow automation devices to be directly integrated with

high-level manufacturing systems and business supply chains.

The vision of the application of DPWS will be a design of control systems into the distributed application paradigms, and ubiquitous computing environments to enable flexible, reusable, reconfigurable manufacturing systems

based on self-reliant, interconnected smart embedded devices [17].

Using the SOA enablement of factory elements these challenges can be addressed potentially in DBE architectures. At Loughborough we have started this investigation in terms of our PDE Tools project. This will now be discussed.

B. PDE Tools

An emerging trend to adopt the Web services approach to the automation device as proposed in [15, 16, and 17] enables the creation of the DBE system. Its concept of WSDPWS is to include the plant activity into the enterprise to facilitate work synchronization between the high and low level applications.

This is a conceptual idea to provide the capability to develop distributed applications in diverse real-time computing environments in many business and automation disciplines including the automotive industry, building automation industry, to name a few.

As we have mentioned DPWS in section 3.1 and it is being developed as the main tool to implement the next generation of factory automation systems.

However to date the research into DPWS is yet to yield any significant implementations. In order to advance into this area, initial investigations at Loughborough into the

leverage of SOA within manufacturing process have been completed.

The PDE (Process Definition Editor) tool developed at Loughborough University [18] contributes a significant role in the design of control application and visualisation at run/design time to ease the design and (re) configuration of the manufacturing process. Currently the tool includes the ability to communicate with factory level elements through web based technologies bridged to PLC's. The further development on this tool will be to support its applications using DPWS.

PDE architecture consists of a library of device profiles and a storage service that stores the data from previous executions of the elements.

The elements are visualised in the PDE tool and these visualisations are linked to the device profiles and saved automated layouts that consist of multiple elements. At the moment the data into the PDE is taken from the PLC controllers of the automation line, an overview of the PDE test bed can be seen in Figure 1.

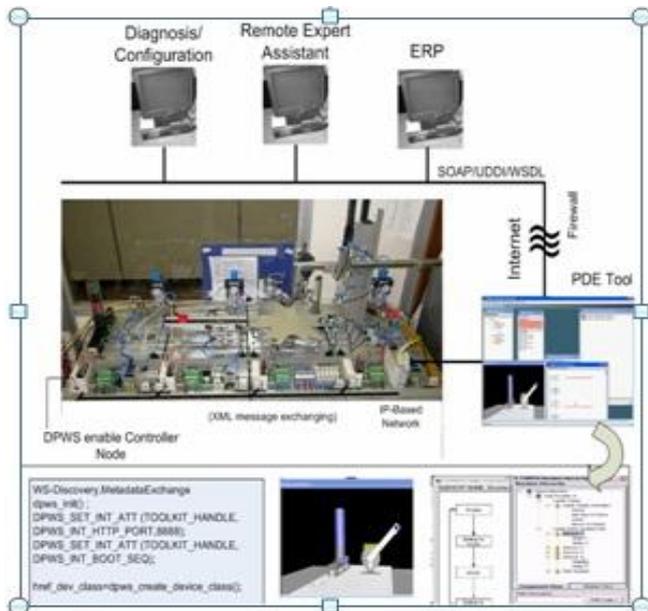


Fig 1: Overview of PDE Test Bed.

DPWS support will present the software with a standardized SOA architecture and will enhance its possible interfaces to higher systems. Currently a SOA can be seen to exist in that common interfaces to the PLC are exposed making the invocation of the PLC possible as a distributed service.

Adoption of the PDE tool to interface directly with devices instead of the PLC controllers will further enhance the model. This could yield new applications in SOA that are already present in PDE tool such as the process of remote configuration based on distributed web technology. The main challenge of DPWS enabling the devices is the hardware needs, at Loughborough we are developing hardware to support DPWS in the form of RTOS that are supporting DPWS and gSOAP [19] linked to the elements inspired by ARM MCU processors [20], as the model presented in figure 2.

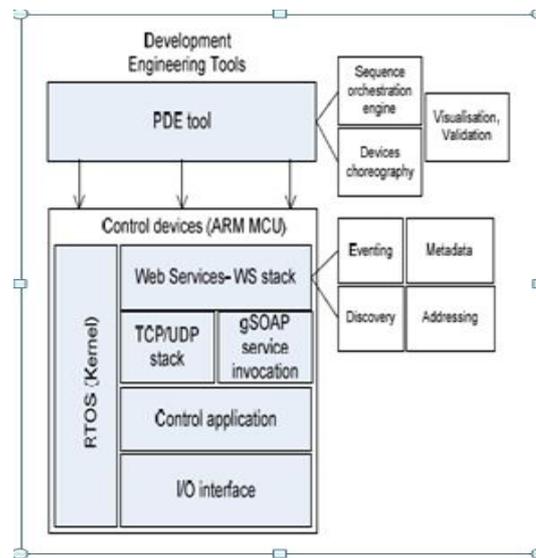


Fig 2 Web services enabling device implementation model

The approach in figure 2 will give greater control over the factory floor devices. However to date the results of our investigation that have been set around new applications developed using the PDE Tool approach, applications that can be seen as an initial insight into the potential of the further SOA enablement of the test bed.

In order to aid explanation, the applications can be divided up into three main categories. The first being visualisation and reconfiguration applications, the second higher level enterprise integration and third lower level integration.

The makeup of each one of these layers will be now discussed focusing on the impacts in terms of new users and applications in this next generation SOA enabled DBE's.

IV. NEW APPLICATIONS FOR THE DBE.

The results of the PDE tool investigation is the creation of new sets of applications that can be used over a distributed automation platform. These will now be discussed focusing on the impacts in terms of new users and applications in this next generation SOA enabled DBE's.

A. Visualisation and Reconfiguration.

From the PDE Tool work a library of components (PLC, Embedded processor board) is being created as a standard catalogue for machine builders using the Process Definition Editor (PDE) tool developed at LU. Components in the library are defined with their logic state operations.

This tool allows a system designer to compose and validate the system using the most appropriate system views (such as State Transition Diagrams, process flow diagrams and a 3-D modeling environment). The tool includes simulation and debug facilities to test and validate the system before and after implementation.

These Libraries are linked to Virtual Reality Modeling Language [21] models of components and enable visualization of the manufacturing process remotely, or repeat visualization of saved past processors. Both the recreation of these runs and also the data saved during this process introduces new capabilities and possible applications to the PDE:

Using standardised Web Services the use of libraries may be enhanced using WS-Resource, although at the moment DPWS is not being developed to support this [22]. The visualisation process could also be enhanced with the WS being configured to collate and broadcast diagnostic and process information to assist support personnel. In fact new support DBE's could be formed around significant manufacturing equipment lines.

As groups of equipment specialists and maintenance engineers who could use this to monitor, analyze, and document the information to schedule maintenance work or examine machine failure modes for proactive and reactive maintenance, accountants who may examine machine throughput/ uptime etc to predict profit, or machine builders/technology providers providing DBE based remote expert assistance.

B. Higher Level Integration.

The adoption of the next generation of applications that the SOA enablement of factory elements offers is dependent on efficient links to higher level business systems. The architecture to enable these links is subject to work in projects such as SOCRADES and has to be tailored for the specific requirements of the factory floor elements. These functions need to take into account both the requirements of the higher level application such as the ERP system whilst protecting the security and safety of the control of the lower level device.

PDE supports the Device Protocol advertising that is apart of the DPWS discovery function using WS-Discovery. Web services discovery message over the network. to accept this device into the system. The device description service is used as the Meta-data file of the device defined in the DPWS protocol stack. This file describes the function/application services the device provides, including the part name, serial number, programme version and etc. This mechanism is handled with the resource management programme tool (PDE).

Work in the PDE tool project has highlighted that the demands of the higher level system need to be catered for in a different way to usual enterprise computing models. For example the typical model of business integration is in the area of supply chain management or automated eCommerce. Here the Larger suppliers require that the SMEs integrate with their ERP type system in order for them to become a trading partner [24]. Here the SME gives certain control of data query and even input into their enterprise type system to the larger partner. Logically the model when expanded to factory floor elements will be repeated. With the business's being able to interrogate data on production machinery and even influence production metrics.

However this control when dealing with lower level machinery has not only business critical issues such as the input of wrong data as is contended within the ERP model, but also the data on live machinery has safety aspects to take into account. For example questions need to be raised about how partner control of live production lines and machinery can be achieved without risk to the equipment and employees using it.

Within the PDE project the virtualisation of the process is achieved so the partners can influence production on a conceptual level. This can be based on simulated live

run data or past data of production runs, and could be linked to higher level business process language such as BPEL. It is likely that the higher level systems will need to integrate with this to enable the control over the lower level machinery. It is therefore likely that this virtualisation will act as a

bridge to either aid understanding from the higher level of a production process, or for requests of changes from a higher level to be represented to an appropriate decision point at the local factory for implementation. Whilst this may mirror current distributed SOA policy and security systems in distributed enterprise research, it is likely this challenge will require these frameworks to adapt to support the new requirement of the manufacturers.

Therefore the PLC is wrapped as a service and the actions of the elements modeled virtually in the PDE system. Overall supply chain management in DBE will benefit from the advanced pervasiveness of the applications that can use data derived from elements on the factory floor. The DBS using SOA enabled elements would be able to report on failures of machinery in real time as with other production statistics. For the DBE this data will influence the way existing and emerging applications function between the higher levels to the lower.

C. Lower Level Integration.

Lower level integration into DBE's is dependent on understandings of lower level process. We are developing algorithms to model the functions at a lower level, this is represented in the behavior of the virtual elements in PDE and would be tied to real element configuration in the device level SOA implementation. The algorithms that we use to model the lower level choreography should be mapped to the higher level workflows in the DBE on integration.

In addition to the implementation of the new installation device for the reconfigured process, the web services facilitates the device discovery service by broadcasting the Web services discovery message over the network (WS discovery) to accept this device into the system. The device description service is used as the Meta-data file of the device defined in the DPWS protocol stack. This file describes the function/ application services the device provides, including the part name, serial number, programme version and etc. This mechanism is handled with the resource management programme tool (PDE).

Based on this approach, dynamic changes of the process have been optimized through the loosely coupled hardware and the web services ontology. In future DBE's these ontology's and choreography models will be as significant as workflows are in current WS enabled distributed DBE's. As in order for the DBE to function correctly at a distributed application level some knowledge of the types of ontology's used will be needed to improve interfaces with the factory floor elements.

V. TRANSFORMING THE DBE.

A. Impact of the New Applications

This paper has only scratched the surface in suggesting the new applications this integration can present to the future DBE. As the distributed enterprise evolves to support the new information yielded from the factory floor

to current DBE models will change and new applications emerge. Research into the PDE project into the visualisation of data from the automated manufacturing process and also integration of lower and higher level devices in the model has suggested significant new areas of development for the DBE.

These three main areas that the paper has highlighted will cover the main area of application development. Within this greater potential for management of lower level devices suggests that new controls at middleware level to bridge the SME and larger customer need to be enabled. This would impact current DBE that support automated supply chain management.

Another area that can be seen as holding huge potential for further development in the DBE is in visualisation of processes and process data. As the information and the knowledge of the manufacturing process can be seen to flow well into applications in this business ecosystem environment. Therefore, in this context, the DBE system is able to be greatly enhanced with Remote expert assistance, Decision support system, Machine design-synthesis and validation. The quick response to the external and internal of business changes can be deployed in the shorter time since the system applications are well defined and integrated with

the standard and platform neutral of web services architectures spread throughout the business ecosystem. To date many DBE's are motivated by higher level processes such as in supply chain based DBE's. The new data from automation system will require a greater management and linkage between higher level applications and lower level processors. Examples of new areas of middleware development and support can be seen to have the potential to develop. Particularly around the needs of safety and control in DBE's that link into lower levels.

Therefore, in the optimisation of the DBE system, it is foreseen by this research such that the inclusion of services from the manufacturing process will have direct impact on the overall development of future DBEs. The new application in service oriented architecture has focused on enhancing the agility, flexibility, as well as robustness of the manufacturing process to support the manufacturing's lifecycles. Also the implementation of the DPWS- Web services on control devices paves the way to enable the creation of a neutral platform where devices from different vendors can interoperate via the XML message passing; hence the end-users are not tied to specific vendors. In addition, system

flexibility is achieved through plug-and-play discovery (UPnP) and binding mechanisms via the use of stub and proxy code for remote procedure calls with the development of DPWS device stack.

VI. FUTURE WORK.

The next stage of the project is to transfer the proposed framework to enable service oriented ecosystem through UPnP and Service orchestration. This will involve in implementing the proof-of-concept DPWS on embedded devices with developed suite of engineering tools. Some of the work has been initiated and done in integrating RTOS with the DPWS on the ARM9 device. Meanwhile, the work will be carrying on enhancing the PDE tool for managing

complexity of the control application and business application integration in a new DBE system. This work will be done on an industry-standard test rig, supported by Ford Motor Company. The research will be investigated in the performance and reliability of control devices in an event-driven control soft real-time environment.

VII. CONCLUSION

This paper has reviewed and proposed key areas of development for a new generation of DBEs for SMEs using the PDE project as a testbed.

The key characteristic of a DBE is to find the optimum solution for the enterprise among business alliances for the skills and knowledge sharing.

Therefore the spread of the DBE is concerned in the integration of two main areas- technical and organisation to support the growing requirement on agility and responsiveness business approach.

The paper has addressed this proposing three application layers of the impact of the factory floor automation system on the DBE.

These layers will be advanced with the deployment of Web services technology and DPWS stack at device level.

The PDE methodology has given an insight into applying web services to automation systems.

The key factor being the implementation of Web services on control devices aims to enable the creation of a neutral and open platform where devices from different vendors can be mixed in the system.

The benefits of service orientation are clearly conveyed all the way to the device level, facilitating the discovery and composition of applications by reconfiguration at the higher level to redefine the combination of provided service of automation devices in the manufacturing process.

In addition, the paper has stressed in dynamic self configuration of smart embedded devices using loosely coupled services provides significant advantages for highly dynamic and ad hoc distributed applications to achieve system flexibility through plug-and-play discovery and binding in support of SOA-Web services.

REFERENCES

- [1] D&ME Pacific North West National Laboratory, "Agile Manufacturing", <http://www.technet.pnl.gov/dme/agile/index.stm> last accessed 1st January 2008.
- [2] Leslie J Lee, "A Next Generation Manufacturing Control System", *PhD dissertation, MSI Research Institute, Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, 2003.*
- [3] SOCRADES project homepage www.socrades.eu/ last accessed 1st January 2008.
- [4] SODA project homepage www.soda-itea.org last accessed 1st January 2008.
- [5] SIRENA project homepage www.sirena-itea.org last accessed 1st January 2008.
- [6] EU Digital Ecosystems Website: www.digital-ecosystems.org last accessed 1st January 2008.
- [7] Clifton Burton, "Information and Communications Technologies: Are They the Key to Viable Business

- Development Service for MSMEs? MICRO ENTERPRISE BEST PRACTICES” *Development Alternative, Inc., 1999.*
- [8] Lavrin A., Pupkov K., Zelko M., “Business Information and Knowledge Services for SMEs”, *IDIMT – 2004. C. Budejovice, Proceedings, edited by Ch. Hofer, G. Chroust. Schriftenreihe Informatik Band 12, Universitätsverlag Rudolf Trauner, 2004, ISBN 3-85487-665-3, pp.269 – 282.*
- [9] Nachiava, F. with contribution of E. Chiozza, H. Ithonen, M. Manzoni, F. Cuningham (2004).”Towards Network of Digital Business Ecosystems Fostering the Local Development; Bruxelles”, 2002, available from <http://www.opencontent.org/open-publ/>.
- [10] European Commission research website “Strong political and industrial consensus on future of European manufacture” http://ec.europa.eu/research/industrial_technologies/articles/article_377_en.html last accessed 1 January 2008.
- [11] SAP Netweaver homepage: <http://www.sap.com/platform/netweaver/index.epx> last accessed 1 January 2008.
- [12] T Kirkham et al “A business service network to aid collaboration between small to medium enterprises”. *Proc. 8th IEEE International Conference on and Enterprise Computing, E-Commerce, and E Services.*
- [13] R. Harrison, A. W. Colombo, A. A. West and S. M. Lee: “Reconfigurable Modular Automation Systems for Automotive Power-Train Manufacture”, *International Journal of Flexible Manufacturing Systems, Springer-Verlag, London. 2007.*
- [14] François Jammes, Harm Smit, “Service-Oriented Paradigms in Industrial Automation”, Schneider Electric Corporate Research, France.
- [15] Robert van Engelen, “Code Generation Techniques for Developing Light-Weight XML Web Services for Embedded Devices”, the ACMSIGAPP Conference, 2004 (Embedded System Track).
- [16] G.B. Machado, F. Siqueira, R. Mittmann, C Augusto, “Embedded System Integration Using Web Services”, IEEE, the international conference on networking (ICNICONSMCL’06).
- [17] R. Harrison, A. W. Colombo: “Collaborative Automation: “From Rigid Coupling towards Dynamic Reconfigurable Production Systems”, Proc. of the 16th IFAC World Congress, Prague, Czech Republic. July 2004.
- [18] P Phaithoonbuathong, R Harrison, C S Mcleod, “A Web Services Based Automation Paradigm for Agile Manufacturing”, 4th international conference on responsive manufacturing, ICRM 2007.
- [19] Robert A. van Engelen and Kyle Gallivan, “The gSOAP Toolkit for Web Services and Peer-To-Peer Computing Networks” Proceedings of the 2nd IEEE International Symposium on Cluster Computing and the Grid (CCGrid2002), pages 128-135, May 21-24, 2002, Berlin, Germany.
- [20] ARM homepage www.arm.com/
- [21] VRML <http://www.w3.org/Markup/VRML/>
- [22] T Kirkham et al “An Agent Driven Assessment of the Potential of DPWS Enabled Automated Manufacturing” submitted IFAC 08.
- [23] Windows Rally technology homepage <http://www.microsoft.com/whdc/rally/default.mspx> last accessed 1st January 2008.
- [24] TS Dillon, Chen Wu, Chang EJ Reference Architectural Styles for Service-Oriented Computing, Keynote, IFIP Networked and Parallel Computing, Dalian, China, LNCS 4672, pp 543-555.
- [25] P Phaithoonbuathong, T Kirkham, C S Melod, M Capers, R Harrison, R P Monfared Reference Adding Factory Floor Automation to Digital Ecosystems; tools, technology and transformation, Loughborough University, Loughborough, Leicestershire, UK, LE11 3TU.

Heat Transfer Analysis of Helical Strip Insert with Regularly Spaced Cut Sections Placed Inside a Circular Pipe

Prof. Naresh B. Dhamane¹, Prof. Mathew V. Karvinkoppa²,
Prof. Murtuza S. Dholkhwala³

(Faculty Mechanical Engg. Department, M.M.I.T., Lohgaon, University Of Pune, India)

(Faculty, Mechanical Engg. Department, JSPM's Imperial College of Engg. & Research, University Of Pune, India)

ABSTRACT : This paper presents an experimental study of heat transfer and friction characteristics in turbulent flow generated by a helical strip inserts with regularly spaced cut passages, placed inside a circular pipe across the test section. The experiments were conducted for water flow rates in the range of Re 5000 to Re 30000. For the experiment three different types of helical strips with helix angles of 30° , 45° and 60° were used. Experimental results show that, the use of a helical strip inserted inside a circular tube results into an enhancement of heat transfer rate as they cause the turbulence in the flow with swirling moment. The local heat transfer coefficients were found to be increasing to very high values along the downstream of the helical strip, and then decreasing with the distance (x/L). The increase of heat transfer was found to be dependent on the Reynolds number in typical case. The effect of the number of the helical channels, and helix angle, on heat transfer was minute. It is found that using the helical tape can help to increase the heat transfer rate up to 20% depending on Re at constant pumping power. Enhancement efficiency decreases with increasing Reynolds number.

Keywords: Enhancement, Helical strip, Heat transfer, Reynold's number, Turbulent flow.

I. Introduction

Turbulent flow has been used in a wide range of applications from various engineering areas such as chemical and mechanical mixing, combustion chambers, turbo-machinery to pollution control devices. It is commonly known that the swirling moment enhances the heat transfer mainly due to the increased velocity in the swirl tube and the circulation of the fluid by centrifugal convection because the low density of the warmer fluid at the pipe wall is displaced into the cooler stream in the central region by centripetal force. Number of attempts have been made to reduce the size and cost of heat exchangers. For this purpose, first a study of the heat transfer coefficient and friction factor for the system under investigation must be carried out. A lot of techniques have been proposed by many investigators for the improvement of heat transfer. Smithberg and Landis [1] reported friction and forced convection heat transfer characteristics in tubes fitted with twisted tape swirl generators, and presented a correlation for predicting Nusselt number and friction factor. Date and Singham [2], and Date [3] reported the prediction of fully developed flow in a tube containing a twisted tape. Hong and Bergles [4] correlated heat transfer

and pressure drop data for twisted tape inserts for uniform wall temperature conditions using water as working fluids in laminar flow. Eiamsa-ard, and Promvonge [5] reported the enhancement of the heat transfer in a tube with regularly spaced and full-length helical tape swirl generators, and concluded that the full-length helical tape with rod provide the highest heat transfer rate about 10% better than without rod.

To increase heat transfer coefficient, rotating or secondary flow on axial forced flow, such as inlet vortex generators, twisted-tapes and axial core inserts are being used. Some of these methods generate swirl flow continuously along the entire length of test section; while others are fixed at the inlet, and decay of swirl along the tube is permitted. These techniques can increase heat transfer coefficient due to related flow patterns, giving the facility of designing smaller exchanger or of upgrading existing exchangers [6,7]. In other applications, effective heat removal may be desirable to prevent excessive temperatures when a certain amount of energy is to be dissipated over a limited area. In the swirling flows pressure loss occurs so they bring additional pumping power and construction costs, and total cost of a heat exchanger increases [8,9]. Bergles [10] cited nearly 250 publications on swirl flows until of the end of 1983. Approximately 140 of these articles were related to heat transfer enhancement. F. Change and Dhir [11] successfully used tangential injection, to obtain swirl flow for the enhancement of critical heat flux. They obtained an average enhancement of 35 to 40 percent in heat transfer on a constant pumping power basis. Another technique of enhancing heat transfer coefficient is the use of the vertical flow. A detailed description of this kind of flow and its potential in heat transfer enhancement were given by Razgaitis and Holman [12]. According to them, one promising technique for the augmentation of convective heat transfer is the use of confined swirl flow, particularly of the decaying vortex type. Kreith and Margolis [13] suggested that swirl flow obtained by tangentially injection of some part of the fluid upon the axial channel flow can lead to an enhancement in heat transfer. But they did not carry out any experiment to support their proposal. Although extensive research has already been done on enhancing heat transfer in tubes inserting a swirl element along the axis or at the entrance, the past effort has focused on the effects of geometrical properties of the swirling elements on heat transfer and frictional losses. In this study, flow ratio (the ratio of swirling flow to non-swirling flow) is considered on account of heat transfer enhancement and pressure drop in turbulent flow. The objective of this research is therefore to investigate the effect of inserting a

swirling element inside the circular tube—on heat transfer and pressure drop.

II. Experimental setup

The flow loop used for this study is shown schematically in Fig. 1.1. The system is consisted of mainly a pump, visual flowmeters, a test section, pressure transmitters, a reservoir tank, data logger and thermocouples. The experimental set-up comprises of two parts: the swirl generator and the test section. The test section is made of smooth stainless steel pipe with the dimensions of 1000 mm length and 20 mm inner diameter. The details of the helical tape design were also shown in Figs. 1.2. The inner diameter and the length of the helical tape are $d = 13$ mm and $l = 100$ mm, respectively. Helical channel is milled on this tape with various angles ($\alpha = 30^\circ, 45^\circ, 60^\circ$) and numbers ($N_h = 1, 2$). Also the rectangular passages at having distance of pitch 10mm have been cut on strip. The depth of the channel is 3.5 mm. The helical tape is tight fitted to the main pipe. The total flow rate is measured with a main flow meter, thereafter the flow is divided into two ways:

One to the inner part of this helical tape and the other through the helical channel. Eventually, the total flow at the exit of the helical strip includes the flow through the inner and outer parts of the helical strip. The test section was heated with current provided from a 15 kW DC power supply. Consequently, the pipe is heated uniformly and in order to prevent voltage fluctuations, a voltage regulator was employed. Glass wool is used as insulation on the outer surface of the test section, in order to reduce heat loss. The entrance and exit temperatures of the air and the wall temperatures were measured with copper-constant thermocouples at ten axial stations, located just after the helical tape at the distances of $x/L = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$ from the inlet to the end of the test section. The thermocouples were isolated with a very thin sheet of mica between the thermocouple and the tube surface so as not to be effected from electricity. All temperature data from the test section were recorded via a data acquisition module. The bulk temperatures of the fluid at the inlet were measured with the two thermocouples. The local bulk temperatures of the water were calculated

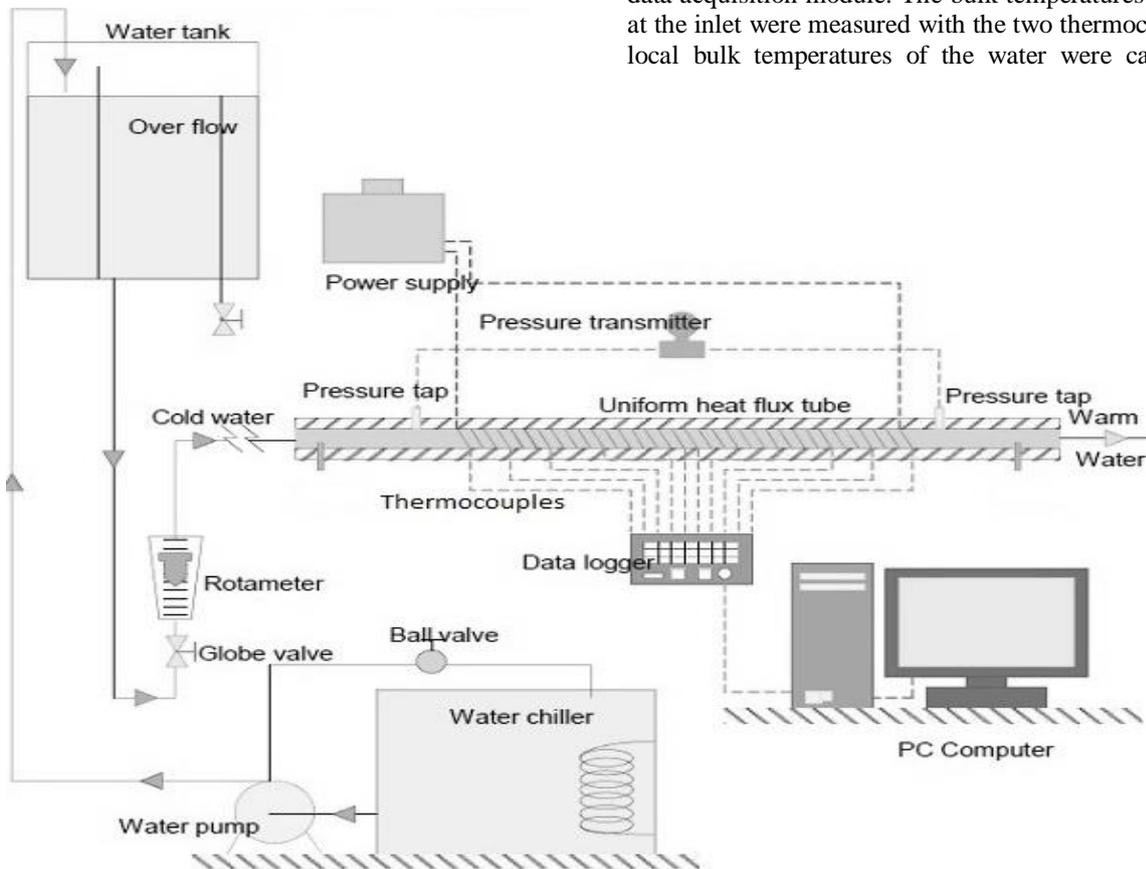


Fig. 1.1. Experimental setup diagram of the flow loops

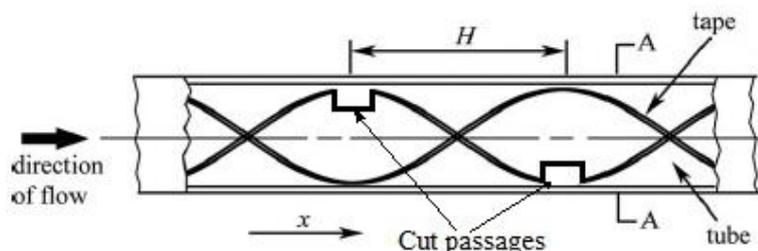


Fig.1.2 Construction of helical strip with regularly cut passages

assuming a liner relationship between the inlet and exit bulk temperatures. This linearity results from a constant heat flux condition with negligible heat transfer along the pipe length [14,15]. Two pressure taps were located on the test pipe to measure the differential pressure drop along the length of the pipe while the other two pressure taps were located on the helical strip.

III. Experimental uncertainty

In this study, the estimations method of Moffat [16] was used. The effect of the uncertainty in a single measurement on the calculated result, if only that one measurement were in error would be as shown in equation (1),

$$\delta R_{x_i} = \frac{\partial R}{\partial X_i} \delta X_i \tag{1}$$

When several independent variables are used in the function R, the individual terms are combined by root-sum-square method,

$$\delta R = \left\{ \sum_{i=1}^N \left(\frac{\partial R}{\partial X_i} \delta X_i \right)^2 \right\}^{1/2} \tag{2}$$

Considering the relative errors in the individual factors denoted by x_n , error estimation is done with the following equation,

$$\epsilon = \left[(x_1)^2 + (x_2)^2 + \dots + (x_n)^2 \right]^{1/2} \tag{3}$$

Nusselt number uncertainties can be calculated by combinations of Eqs. (4) and (5),

$$Nu = \frac{hD}{k} \tag{4}$$

$$\epsilon_{Nu} = \left\{ \left(\frac{\partial Nu}{\partial h} \epsilon_h \right)^2 + \left(\frac{\partial Nu}{\partial D} \epsilon_D \right)^2 + \left(\frac{\partial Nu}{\partial k} \epsilon_k \right)^2 \right\}^{1/2} \tag{5}$$

$$\frac{\epsilon_{Nu}}{Nu} = \left\{ \left(\frac{\epsilon_h}{h} \right)^2 + \left(\frac{\epsilon_D}{D} \right)^2 + \left(\frac{\epsilon_k}{k} \right)^2 \right\}^{1/2} \tag{6}$$

The individual contributions to the uncertainties of the nondimensional parameters for each of the measured physical properties are summarized in Table 1. Maximum values of uncertainty calculations for Re , Nu and f are 7.5%, 10.4% and 13.4%, respectively.

IV. Heat transfer calculations

In order to determine the heat losses, the test set-up was calibrated under no-flow conditions. The calibration curve was approximated by,

$$Q_{loss} = 0.345L (\bar{T}_w - T_a) \tag{7}$$

where L , represents the length of test pipe, \bar{T}_w is average wall temperature and T_a is the average ambient temperature.

The average heat flux from the tube wall to the fluid is defined in terms of the nominal inside surface area,

$$Q = hA(\bar{T}_w - T_b) \tag{8}$$

It took approximately two hours to obtain a steady state for each run. Experiments were conducted for different values of heat input to check the reproducibility of the results.

Table 1 : Values of uncertainties for some variables

Variables	Uncertainty
Density of water, ρ	1.6
Specific heat (Water), C_p	3.2
Thermal conductivity of water, k	2.4
Dynamic viscosity (Water), μ	2.9
Pressure drop, ΔP	4.2
Hydraulic diameter, D_h	1.3
Water flow rate, \dot{m}	4.6

Due to a small temperature gradient along the tube, the axial heat flux was essentially constant; hence it was possible to utilize the average heat flux to compute local heat transfer coefficients. The heat rate was assumed to be uniform. The thermal conductivity of test tube is a strong function of temperature. Its functional dependence on temperature was obtained from the work of Hogan [17]. The bulk temperature at each position was computed assuming a linear variation along the heated length, and heat transfer coefficient was then found from the following equation,

$$h_{x,i} = \frac{\dot{q}}{T_{w,i} - T_{b,i}} \tag{9}$$

where, $(h_{x,i})$ is the local heat transfer coefficient, found as $\dot{q} = (Q_{net}/A)$ is the local heat transfer rate per unit from the wall to water, $T_{(w,i)}$ is the local wall temperature of the tube and, $T_{(b,i)}$ is the local bulk mean water temperature. The inside tube wall temperature was obtained by correcting the measured outside wall temperature by solving the cylindrical heat conduction equation from measured outside wall temperature; the one dimensional steady-state heat conduction equation with variable thermal conductivity was solved numerically. The local liquid bulk temperature was calculated from the heat balance and local Nusselt number was calculated as,

$$Nu_{x,i} = \frac{h_i D_h}{k} \tag{10}$$

$$Re = \frac{u D_h}{\nu} \tag{11}$$

The local values for Nu , Pr and Re were calculated based on water properties corresponding to the bulk fluid temperature, and the average values were obtained by numerical integration. In Eq. (11) Re is defined based on the total flow in the test section. In the experimental results, normalized Nusselt number is defined as follow,

$$Nu^* = \frac{Nu_s}{Nu_o} \tag{12}$$

where Nu_s is Nusselt number for swirl flow and Nu_o is the Nusselt number for fully developed axial flow in the tube without helical strip calculated according to Eq. (10). In this work, swirl intensity is expressed as the ratio of tangential momentum of the helical fluid to the axial momentum of the total flow and this equation expressed by Guo and Dhir [18] as,

$$\frac{M_h}{M_T} = \frac{\dot{m}_h^2 \frac{A}{A_h}}{\dot{m}_T^2} \tag{13}$$

where \dot{m}_h is the flow rate of the water through the helical channel and \dot{m}_T is the total mass flow rate of the water through the test section. A and A_h are the cross-sectional area of the test section and total area of the helical channel, respectively. Eq. (13), can be rearranged for helical strip as,

$$\frac{M_h}{M_T} = \frac{\dot{m}_h^2}{\dot{m}_T^2} \cdot \left(\frac{W.H}{\pi D^2/4}\right) \cdot \frac{\sin \alpha}{N_h} \quad (14)$$

where N_h is the number of channel on the helical tape, W and H are the helical channel width and height, respectively. It is important to compare the experimental data obtained for Nusselt number in fully developed axial flow with the correlations from the literature. Nusselt numbers calculated from the experimental data were compared with the correlation obtained by Petukhov [19],

$$Nu_0 = \frac{(f/8)Re Pr^{0.4}}{1.07 + 12.7 \sqrt{f/8(Pr^{2/3}-1)}} \left(\frac{\mu_b}{\mu_e}\right)^{0.11} \quad (15)$$

(for $T_w > T_b$)

f , is friction factor and for smooth tubes it is given as ,

$$f = (1.82 \log_{10}(Re) - 1.64)^{-2} \quad (16)$$

So as to determine the friction factor for both test section and helical strip, pressure drop across the helical channel and across the test section are measured separately. The friction factor is then obtained by the following equation,

$$f = \frac{\Delta P}{\left(\frac{L}{D_h}\right)\rho \frac{v^2}{2}} \quad (17)$$

V. Experimental results and discussion

Experimentally determined Nusselt values for smooth tube (without helical strip) are compared with Dittus–Boelter correlation [20,21], Petukov correlation [21], and Sieder and Tate correlation [21] for turbulent flow, in Fig. 3. It is seen that the experimental results of the present work are in good agreement with the aforementioned studies.

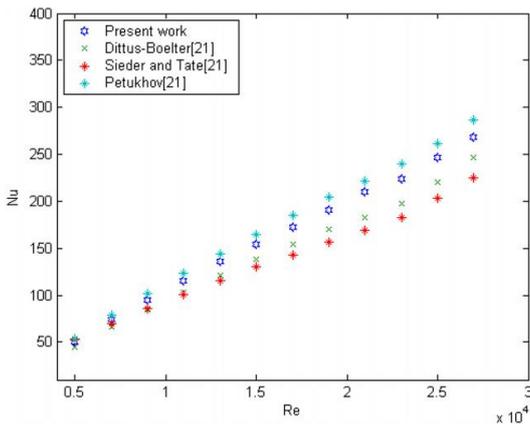


Fig. 3. Data verification of average Nusselt number for the smooth tube.

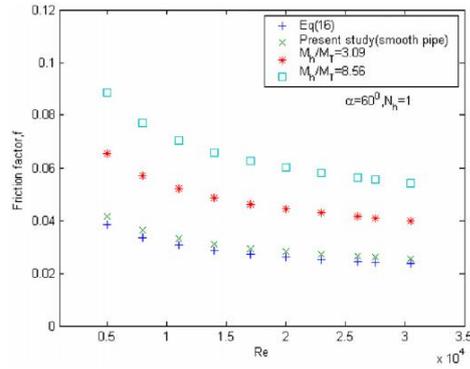


Fig. 4. Friction coefficient variations for the test section with respect to Re.

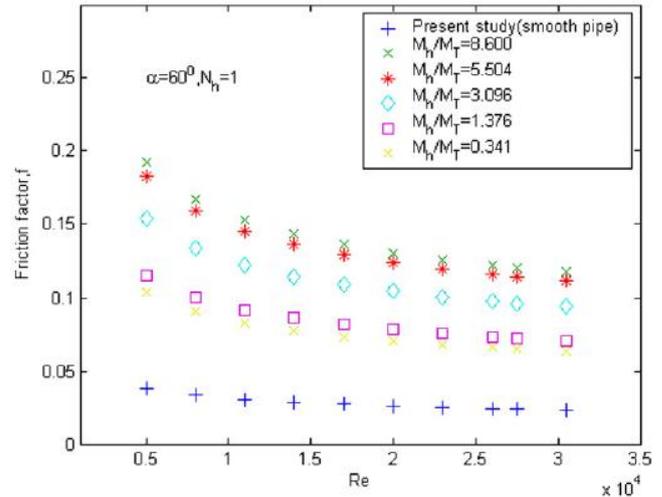


Fig. 5. Friction coefficient variations for the helical channel with respect to Re.

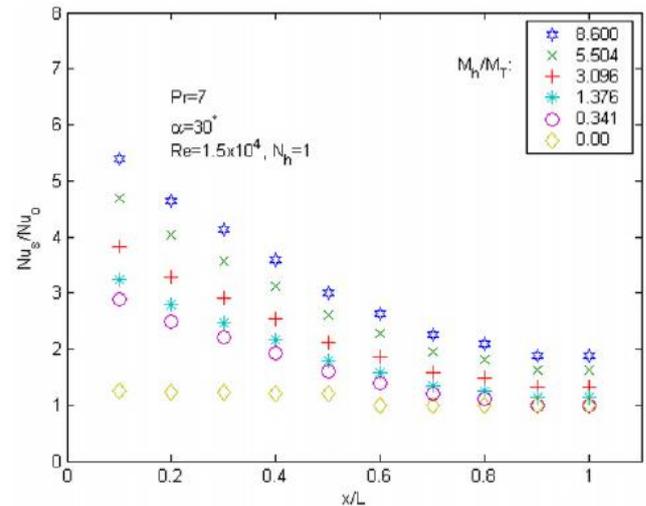


Fig. 6. Normalized Nusselt number distribution along test section for different momentum ratios.

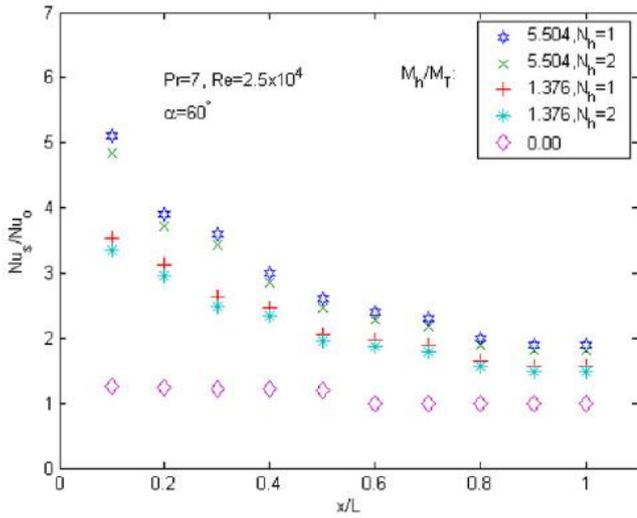


Fig. 7. Effect of the helical channel number on normalized Nusselt.

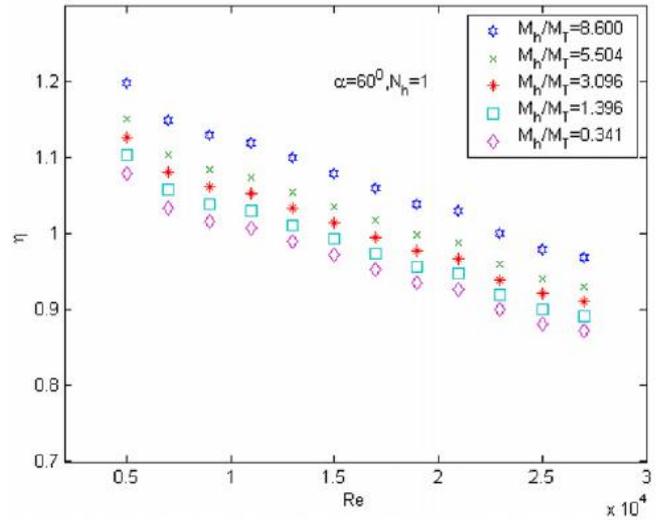


Fig. 10. Variation of heat enhancement efficiency with Reynolds number.

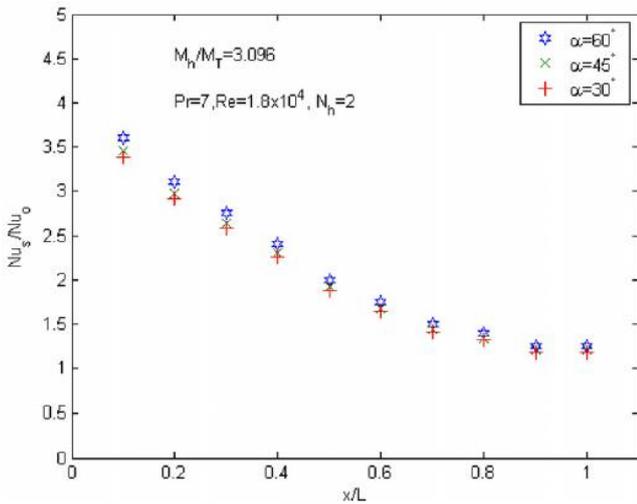


Fig. 8. The effect of helical angle on normalized Nusselt number.

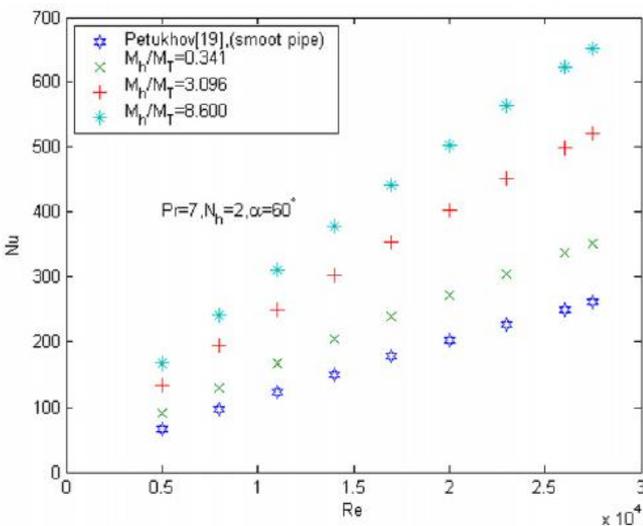


Fig. 9. Variation of average Nusselt number with Re for different momentum ratios.

Fig. 4 shows the comparison of friction factors for the test section determined from the present experimental work and from the correlation of Eq. (16). It is seen that helical strip cause a considerable increase in friction factor. The increment of friction factor for swirl element with $\alpha = 60^\circ$, $N_h = 1$ and $M_h/M_T = 8.56$ is about 2.1 times in comparison with nonswirl flow. The friction factor decreases with increasing Re number and increase with increasing momentum ratio. A similar tendency can be seen in Fig. 5 in which friction across the helical channel is plotted versus Reynolds number. When compared to the friction across the smooth pipe, the friction factor is about 2.5 to 5 times higher than the non-swirl flow depending on the momentum ratio and Reynolds number, as in Fig.5. Ratio of Nusselt number for swirling flow to none swirling represents the heat transfer enhancement. In Figs. 6–8 the heat transfer enhancements for different momentum ratios are plotted as a function of axial distance from the helical tape exit. The increase in heat transfer coefficient was found to be strongly depending on the momentum ratio M_h/M_T . This heat transfer augmentation is entirely due to the thinner thermal and hydrodynamic boundary layers caused by swirling flow. Increases in local velocity with swirl generator may intensify the flow turbulence. It can be seen that local heat transfer coefficient decreases along the tube axis because tangentially momentum decreases due to the viscous dissipation and fluid mixing. So as to illustrate the effect of helical channel number on heat enhancement, normalized local Nusselt number is plotted as a function of axial distance for a constant Prandtl number and for one helical channel ($N_h = 1$) and two helical channels ($N_h = 2$) with 60° angle in Fig. 7. But it is seen from this figure that the main trend for both one and two channels seems to be nearly the same. The enhancement is only a little higher with one helical channel for both momentum ratios. Helical channel number has no significant effect on Nu_s/Nu_0 . It is clear, however, that the increase in heat transfer coefficient at a given location is found to be strongly depended on the momentum ratio M_h/M_T . The magnitude of the heat transfer in swirl flow is much larger than the one observed in the thermally developing region of purely axial flow. The local Nusselt number decrease along the test section with the

increasing axial distance, but at much smaller decreasing rate than for the flow of the smooth tube without a helical strip. This decrease is more at higher momentum ratios. In order to analyze the effect of the helical angle on heat transfer, normalized Nu number is plotted versus x/L for three different helical angles, in Fig. 8. This figure demonstrates that heat transfer enhancement is weakly affected by the helical angle. Average Nusselt number versus Reynolds number is plotted in Fig. 9. As mentioned above for local Nu numbers, increasing momentum ratio causes an increase also in average Nusselt number.

VI. Performance criteria

It is necessary to determine the efficiency of this experimental set-up with and without a swirl flow under the condition of constant pumping power. The heat transfer enhancement efficiency for constant pumping power can be expressed as follows,

$$\eta = \left(\frac{h_s}{h_o}\right)p \quad (18)$$

The heat transfer efficiency versus Reynolds number for various momentum ratios is plotted in Fig. 10. It is seen that the heat transfer efficiency increases with the increasing momentum ratio while decreases with increasing Reynolds number. Enhancement efficiencies varied between 0.97–1.20 for the highest momentum ratio of $M_h/M_T = 8.600$ and, 0.873–1.08 for the lowest momentum ratio of $M_h/M_T = 0.341$, depending on Reynolds number. It is obvious that, for a net energy gain the value of η must be greater than unity. As can be seen in Fig. 10, η is higher than unity for $M_h/M_T = 8.600$ and $Re > 22\,000$. However, for the lowest momentum ratio $M_h/M_T = 0.341$, η is higher than unity only if Re is bigger than 10 000. Consequently, the helical tape will cause an energy loss rather than gain below the mentioned Re values at which η is lower than unity.

VII. Conclusions

For different experimental conditions, the helical strip with regularly cut passages, the heat transfer coefficients were found to be higher than fully developed non-swirling flow. The highest local heat transfer coefficient was obtained at the highest momentum ratio. However, the enhancement in the heat transfer decreases along the tube due to the reduction in swirl flow. An enhancement up to 250% in local heat transfer coefficient was observed in the swirl flow compared to the fully developed axial flow for the same fluid velocity depending on momentum ratio and Re . Helical channel number and helical angle has no significant effect on Nu_s/Nu_0 . The enhancement is a little higher when the helical angle $\alpha = 60^\circ$. For the highest momentum ratio a net energy gain up to 20% was achieved depending on Re number.

References

- [1] E. Smithberg, F. Landis, Friction and forced convection heat transfer characteristics in tubes fitted with twisted tape swirl generators, ASME J. Heat Transfer 2 (1964) 39–49.
- [2] A.W. Date, J.R. Singham, Numerical prediction of friction factor and heat transfer characteristics of fully developed laminar flow in tubes containing twisted tapes, ASME 94 (9) (1972) 54–58.
- [3] A.W. Date, Prediction of fully developed flow in tube containing a twisted tape, Int. J. Heat Mass Transfer 17 (1974) 845–859.
- [4] S.W. Hong, A.E. Bergles, Augmentation of laminar flow heat transfer in tubes by means of twisted-tape inserts, ASME J. Heat Transfer 98 (2) (1976) 251–256.
- [5] S. Eiamsa-ard, P. Promvong, Enhancement of heat transfer in a tube with regularly spaced helical tape swirl generators, Solar Energy 78 (2005) 483–494.
- [6] W.J. Marner, A.E., Bergles, in: 6th Int. Heat Transfer Conf., Toronto, Canada, vol. 2, 1978, pp. 583.
- [7] A.H. Algifri, R.K. Bahardwaj, Y.V.N. Rao, Heat transfer in turbulent decaying swirl flow in a circular pipe, Int. J. Heat Transfer 31 (8) (1988) 1563–1568.
- [8] A. Durmu,s, Heat exchanger and exergy loss in concentric heat exchanger with snail entrance, Int. Commun. Heat Mass Transfer 29 (3) (2002) 303–312.
- [9] M. Yılmaz, O. Comakli, S. Yapici, O.N. Sara, Heat transfer and friction characteristics in decaying swirl flow generated by different radial guide vane swirl generators, Energy Conserv. Manag. 44 (2003) 283–300.
- [10] E. Bergles, Heat transfer enhancement—the encouragement and accommodation of high heat fluxes, J. Heat Transfer Trans. ASME 119 (1997) 8–19.
- [11] F. Change, V.K. Dhir, Turbulent flow field in tangentially injected swirl flow in tubes, Int. J. Heat Fluid Flow 15 (1994) 346–356.
- [12] R. Razgaitis, J.P. Holman, Survey of heat transfer in confined swirl flow, Heat Mass Trans. Process. 2 (1976) 831–866.
- [13] F. Kreith, D.Margolis, Heat transfer and friction in turbulence vortex flow, Appl. Sci. Res. 8(1959) 457–473.
- [14] H. Algifri, R.K. Bhardwaj, Y.V.N. Rao, Heat transfer in turbulent decaying swirl flow in a circular pipe, Int. J. Heat Mass Transfer 31 (8) (1988) 1563–1568.
- [15] O. Kitoh, Axi-symmetric character of turbulent swirling flow in a straight circular pipe, Bull. JSME 27 (226) (1984) 683–690.
- [16] R.J. Moffat, Describing the uncertainties in experimental results, Experimental Thermal Fluid Sci. 1 (1988) 3–17.
- [17] C.L. Hogan, The thermal conductivity of metals at high temperature, PhD dissertation Lehigh University, 1950.
- [18] Z. Guo, V.K. Dhir, Single- and two-phase heat transfer in tangential injection-induced swirl flow, Int. J. Heat Fluid Flow 10 (3) (1989) 203–210.
- [19] B.S. Petukhov, Heat transfer and friction in turbulent pipe flow with variable physical properties, in: Advances in Heat Transfer, vol. 6, Academic Press, San Diego, 1970, pp. 504–564.
- [20] Y.A. Cengel, Heat Transfer: A Practical Approach, second ed., McGraw- Hill, New York, 1998.
- [21] F.P. Incropera, D.P. De Witt, Introduction to Heat Transfer, John Wiley and Sons, New York, 1990, pp. 456–457.

Privacy Preserving Updates for Anonymous and Confidential Databases Using RSA Algorithm

Ishwarya M. V¹, Dr. Ramesh Kumar. K²

¹(Research Scholar, Hindustan University, CSE Department,
Assistant Professor, Sri SaiRam Engineering College, Anna University, India)

²(Department of Information Technology, Hindustan University, India)

ABSTRACT: This paper aims to speed up the implementation of the RSA algorithm during data transmission between different communication networks and Internet, which is calculated to generate the keys by a program prepared in a C # language and then save these values of the keys in the databases created by SQL Server 2008 R2. Privacy is main concern in the present technological phase in the world. Information security has become a critical issue since the information sharing has a common need. Thus privacy is becoming an increasingly important issue in many data mining applications in various fields like medical research, intelligence agencies, hospital records maintenance etc. This paper suggests advancing the existing database systems and increasing the security and efficiency of the systems. This paper proposes a new concept to implement a real world anonymous database which improves the secure efficient system for protection of data, restricting the access to data even by the administrator thus maintaining the secrecy of individual patients. In RSA algorithm identical database must be used in all networks gateways, the creation of the database controlled by a special protocol programmed in a C # language called RSA Handshake Database Protocol, the protocol controls each gateway that runs a RSA-Key Generations Offline according to specific issues and necessities.

Keywords: privacy, database, security, confidentiality, anonymous, cryptography, RSA, RSA-Key Generations Offline, RSA Handshake Database Protocol

I. INTRODUCTION

In today's world databases represent need for increases security. Data in the databases has its own relevant value. For example; medical data collected by over the history of patients over years is an invaluable asset, which needs to be secured and can be used by people in various related areas of work. [1]

Nowadays, privacy accidents have become common problem in the information systems. For example, a hospital may have record of all the patients with various diseases critical and non-critical. If the hospital wishes to reveal the data to any pharmaceutical company or online market services, it should not be able to infer with particularity of patients with those diseases. It can give as a statistical view or just the superficial information such that privacy is not detained.

There are huge numbers of databases that hold numerous confidential information such that people access those data correlating various information from various databases. Access rights for different users must be evaluated and information must be disclosed only to particular extent based on the access rights. Disclosure of confidential information to unauthorized persons may lead to data insecurity leading to dissatisfaction to users. Example privacy accident which occurred so far are numerous. For example, there was a company which sold health products online that also revealed the customer names phone numbers credit card numbers etc on the website. It leads to huge loss of information and breach of privacy. There was another issue when a researcher was enabled to retrieve health records from anonymous databases of insurance claims of employees.

This paper proposes methods to solve the problem of insecurity in database systems by restricting even the administrator from accessing the internal information. The proposed architecture implements the real world anonymous database by implementing the generalization and suppression. It deals with preventing malicious parties and intrusion using encryption and decryption techniques. The efficiency and security of data can be achieved by maintaining single database with specific access rights.

II. RELATED WORKS

In the paper [1] the author suggested paper deals with problems concerning that the users without revealing the contents of tuples and DB, how to preserve data integrity by establishing the anonymity of DB and if the anonymity is authorized then there is a concern of updating the data. It deals with algorithms for database anonymization.

This paper shows how privacy is maintained without disclosing the contents of whole databases and their owner and individual tuples and its owner to each other. The problem is to check whether the database connecting the tuple is still k-anonymous, such that no one can view the actual data from, tuples or database.

In the paper [2] the author suggests this paper is about k- anonymity in wireless sensor networks (WSN). It has a security framework which has two levels of privacy. In this method, some part of the data is encrypted and the rest of the data is generalized.

In the paper [3] k-anonymity concept has been welcomed in many organizations to release micro data without disclosing identity of persons accessing database. The previous k-anonymity techniques implemented in a common database has breached privacy. It leads to loss of information and privacy accidents. From this first we have to introduce a k-join-anonymity, which allows better effective generalization and helps to diminish the loss of data.

In the paper [4] they discuss the relationship between privacy preserving and SMC and problems involved. It reviews definitions and constructions for secure multiparty computation and discusses the issue of efficiency and demonstrates the difficulties involved in constructing highly efficient protocols.

In the paper [5] the anonymization tables were introduced. The issue of releasing tables mainly in relational database consisting of confidential data and how it can be resolved, ensuring personal privacy and also maintaining integrity. One of the techniques proposed in the literature is k-anonymization.

It is k- anonymous if the data for an individual person contained can't be eminent from least of k-1 persons all whose data also shows in the same data release.

Providing username and password and also a random salt value to increase the security and efficiency of the system authentication.

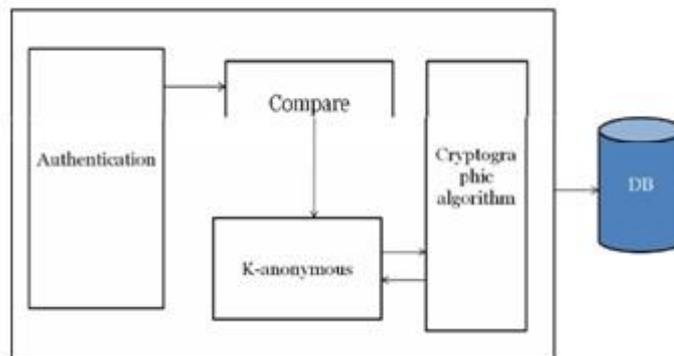


Figure 1: Overview of proposed system

III. BACKGROUND WORK

A. Basics

The existing system has privacy preserving techniques which can be intruded by various sources and privacy is deterred. The anonymization process and cryptographic techniques are used to enhance the efficiency and protection of confidential information in the databases.

Anonymous database features are either suppressed or they are generalized as far as each row remains identical to at least k-1 of the other rows. This is where the database is said to be k-anonymous. The Anonymity thus blocks definite linkages in databases concerning security.

Anonymity gives a definite guarantee that the data is accurately released. The Drawback is the problem of security occurs since the databases are handled by many sources has to be protected.[5]

Cryptographic algorithms and techniques are methods that help to enhance the security of the system and therefore preserve integrity. There are many techniques available in the present technology. The shortcoming of some of the techniques is that they do not provide access rights to the users. Such that for each user has different needs of accessing the database. [6]

The drawback of existing systems is that there is no strong authentication for the systems. If the password for authentication is known to intruder the accounts can't be interrupted and confidential information may be lost. So the authentication should not be able to break even by the administrator, database maintenance, etc.

Moreover, updating of data is a problem in the existing systems due to large amount of data fed in the database. Leads to redundancy and also may lead to loss of valuable information. [6]

Anonymous databases existing do not have proper security and efficiency is low. It can be intruded by various forces. The existing system with cryptographic techniques existing does not grant access rights to the users. [1]

B. Proposed system

The proposed system has features enhanced to existing system. The system is provided with facility for allowing the right users to access into the database by

Proposed system in the figure 1 compares existing data and the updates and make sure there is no

redundancy and helps to analyses the data in database. K-Anonymization allows database to maintain a suppressed and generalized form of data such that data is much secured. The cryptography technique is used to secure the saved data in database safely such that the information is encrypted, stored and can be retrieved and decrypted back to original with specific authorization.

C. Detailed architectural design

The figure 2 shows the flow of steps followed in the system. It starts with authentication of user. Each user is provided with username and password registered in system already. There is a salt value authentication along with password. The authentication user has access to the database and system has particular access rights for each user. The anonymous database suppresses and generalizes the data according to data value.

The database can be accessed by research centres for gathering statistical data regarding particular medicines, the percentage of curable medicines. The internal or private information of the patients are not revealed to the research centre computation. The research people can see the data's send by the database according to its access right. And allocate research peoples to each research data. And forward the data to research people. Here research people can't do any changes or modifications in patient database they only can use the database for reference purpose.

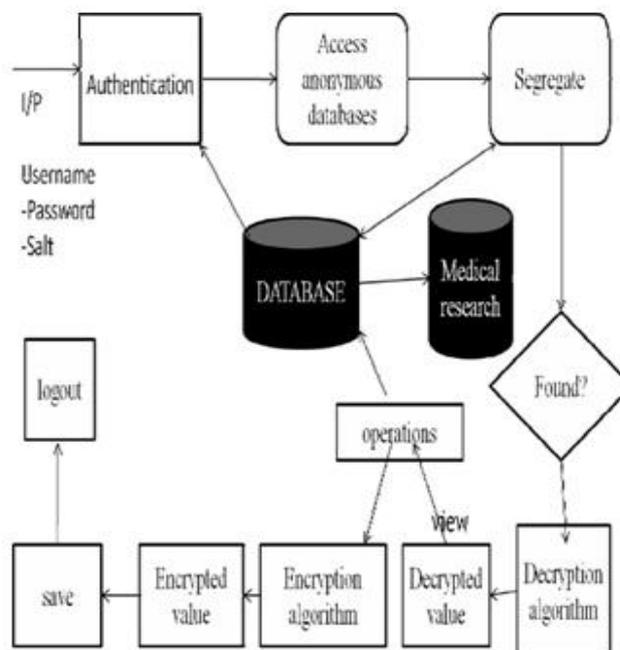


Figure 2: Design view of system.

The authorized database updaters can login into the medical. Here also all the details about the database updater are registered by the admin. And the admin give the authentication details to the particular updater after getting the authentication details, can login to the database and can start the processes.

The doctor and patient enter all details regarding their treatment details in the database in the hospital. These details are not disclosed to the research centre. The data can be encrypted and saved and can be decrypted back to original form when required. [6]

For example, in the proposed system, even the administrator has only restricted access to database, he can't access the internal details of each user, and rather he can find how many users are updated and solve issues regarding users. Individual users are not allowed access to other accounts except their personal record.

Figure 3 shows view of the system for an example patient database where there is a medical database, where the patient registers their details initially. The doctor can view their necessary information about the patient and also fix appointments for patients easily. The patients can in turn fix appointments with their doctor in charge and reduces waiting time for patients due to appointments.

The doctor can in turn update record of patients and their treatments to the patient database. Also the doctor can and their treatments.

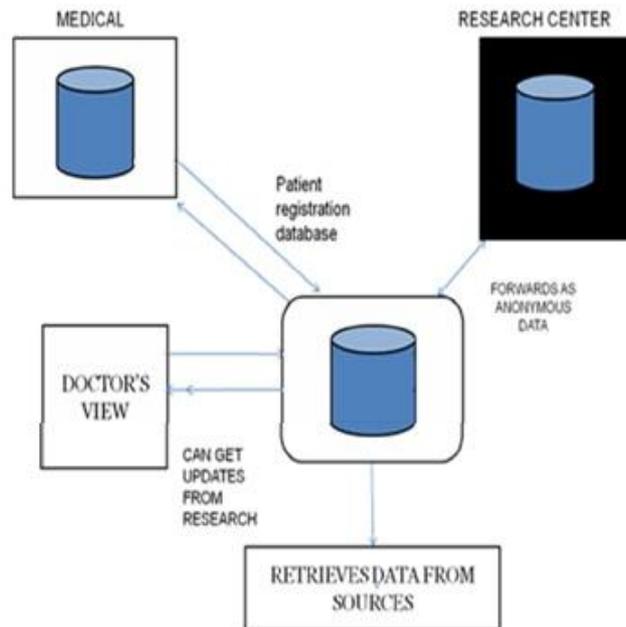


Figure 3: An example diagram of patient database system implementing the idea

The anonymous database can forward information to research centre which has permissions to access the information. The research centre access has its own restrictions for use of the data. They can access only superficial data and whatever data that they have access to. They cannot access the patient details or the particular patient illness.

Figure 4 shows the flow of activities concerning the system. Firstly, the user is authenticated entry into the system. After authentication, according to the user access rights the user can access the anonymous databases in the system. The patient and doctor can perform necessary updates. These data in the anonymous databases can be access by users according to their privileges. The users can access the information from their profile pages. After validating the data and access of data users can close and logout from their account.

D. Implementation of RSA Algorithm

1.1. Offline RSA-Key Generations

In this paper we increased the RSA implementation speed by generated keys offline and stored in different databases before starts using the RSA key pair in encryption/ decryption processes.

RSA-Key Generations Offline is a new software component we developed by using C# language to increases the speed of RSA implementation [NAG 11] [WEL 01], also we need database engine to save the calculated values inside two tables, table one includes the values of p , q , n and $\phi(n)$, and table two includes e and d values.

$$n = p \times q \quad (4)$$

$$\phi(n) = (p - 1)(q - 1) \quad (5)$$

$$e = \text{relatively prime to } \phi(n) \quad (6)$$

$$d = e^{-1} \text{ mod } \phi(n) \quad (7)$$

1.2. Online Encryption and Decryption Processes

In this paper we proposed four security levels each level has own database and consists of many sets, these levels identifiers by possibility of e values and the key length see table 1.

The gateways (users) must select the same security level or change the security level before start the encryption and decryption processes.

We select SQL Server 2008 R2 as database engine for creation the databases and their sets which contents the keys values, also we select SQL Server 2008 R2 to keep our database saves and secure, by encrypted all data without increasing database size or impacting performance and it has Guard against security breaches if backups or disks are lost or stolen.

Table 1. Security Levels

Security Level	Key Length
Low	512 bits
Medium	1024 bits
Medium -High	2048 bits
Security High	4096 bits

In this paper we proposed to use RSA key pair between LAN's / WAN's gateways instead of users. Using of private and public keys between gateways that means the RSA encryption/decryption algorithm now is suitable for large amount of data flow between gateways and this infer of uses the RSA-Key Generations Offline Algorithm, in figure 2.

In this paper we proposed a new protocol called RSA Handshake Database Protocol, this protocol responsible for creation the identical RSA-Key Generations Offline databases in all network gateways and organize database update if require and execute the procedure for each new gateway want to use the RSA-Key Generations Offline database with existing gateways.

The RSA Handshake Database Protocol saves the selected security level (database), which set selected in the security level (Setid), keys indexes and another data in working information table.

The RSA algorithm starts using the data from working information table for encryption/ decryption processes between network gateways.

The RSA Handshake Database Protocol controls all initially processes and any changes in the security levels and key length between the gateways or new gateway would like to join an existing session.

IV. Exchange the keys indexes

In this paper we proposed a new method called Indexes exchange, where we use the Indexes exchange instead of keys exchange between different gateways, example in table 2 explaining how the indexes will be exchanged instead of n, e and d values.

Table 2. Example of use the Indexes Exchange instated of keys exchange

Keys Exchange		Indexes Exchange	
n	160	Nid	00000001
e	7	Eid	000001010010
d	23	Did	110000000001

By using the indexes exchange instead of keys exchange it will be very hard to get the n, e and d values even if you know the indexes of these values.

V. Experiments and results

With using RSA-Key Generations Offline Algorithm and different keys lengths, the decryption processes is 2.5 times faster than online RSA keys generations.

The timings were made on a 2.8GHz Pentium by using the below factors:

- Block size is 2048 bits.
- Different bandwidths:-
 1. 1000 Mbps.
 2. 100 Mbps.
 3. 4 Mbps.

The figure 3 shows the compare between RSA decryption process by using RSA-Key Generations Offline method and online RSA key generation's method, decryption by RSA-Key Generations Offline is faster than using normal RSA key generations.

VI. Conclusion

In this paper, we speedup the RSA algorithm through developed a new generation keys method called RSA-Key Generations Offline to generate and saved all keys values in tables within database.

We proposed four security levels, each level has its own database and numbers of sets, these levels identified according to the e values and keys length, before start using the RSA algorithm between gateways must get a Ready Acknowledgment from RSA Handshake Database protocol, this protocol responsible for creation or update the identical gateways database, level selections (Setid) and establishment the algorithm between gateways.

In this paper we proposed a new method of keys exchange to increase the difficulty for any one knows the exchanged values between gateways, and then try to get the n, e and values, this method we called Indexes exchange, where we exchange the indexes Nid, Eid, Did instead of n, e, d values.

REFERENCES

- [1] [BAH 10] M. Bahadori, M. R. Mali, O. Sarbishei, M. Atarodi and M. Sharifkhani "A novel approach for secure and fast generation of RSA public and private keys on SmartCard" NEWCAS Conference (NEWCAS), 2010 8th IEEE International, 2010, pp. 265-268.
- [2] [BLA 00] S. R. Blackburn and S. D. Galbraith "Certification of secure RSA keys" Electronics Letters, vol. 36, pp. 29-30, 2000.
- [3] [HGE 06] H. Ge and S. R. Tate "Efficient Authenticated Key-Exchange for Devices with a Trusted Manager"
- [4] Information Technology: New Generations, 2006 (ITNG 2006). Third International Conference on, 2006, pp. 198-203.
- [5] [JOS 08] J. Joshi, et al. "Network Security" Morgan Kaufmann, 2008.
- [6] [NAG 11] C. Nagel, B. Evjen, J. Glynn, K. Watson and m. Skinner "Professional C# 2008" Wrox, 2011.
- [7] [REN 05] H. Ren-Junn, S. Feng-Fu, Y. Yi-Shiung and C. Chia-Yao "An efficient decryption method for RSA cryptosystem" Advanced Information Networking and Applications, 2005 (AINA 2005). 19th International Conference on, 2005, pp. 585-590 vol.1.
- [8] [RIV 78] R. L. Rivest, A. Shamir and L. Adleman "A method for obtaining digital signatures and public-key cryptosystems" Communications of the ACM, vol. 21, pp. 120-126, 1978.
- [9] [SEL 89] A. Selby and C. Mitchell "Algorithms for software implementations of RSA" Computers and Digital Techniques, IEE Proceedings E, vol. 136, pp. 166-170, 1989.
- [10] [STA 00] W. Stallings "Network security Essentials: Applications and Standards" Pearson Education India, 2000.
- [11] [STA 03] W. Stallings "Cryptography and network security vol. 2" prentice hall, 2003.
- [12] [STA 95] W. Stallings "Network and internetwork security: principles and practice" Prentice-Hall, Inc., 1995.
- [13] [TIA 06] C. Tianjie and M. Xianping "Collusion Attack on a Server-Aided Unbalanced RSA Key Generation Protocol" Communication Technology, 2006(ICCT 2006). International Conference on, 2006, pp. 1-3.
- [14] [WEL 01] M. Welschenbach "Cryptography in C and C++" Springer-Verlag New York, 2001.

“Effect of Types of Fine Aggregate on Mechanical Properties of Cement Concrete”

Prof. Wakchaure M. R.¹, Er. Shaikh A.P.², Er. Gite B.E.³

(Department of Civil Engineering, Amrutvahini College of Engineering, Sangamner, Maharashtra, India. 422608.)

ABSTRACT: This paper presents the effect of the use of artificial sand as fine aggregate in concrete as substitutes to natural sand. The experimental work is mainly concerned with the study of mechanical properties like compressive strength, split tensile strength and flexural strength of concrete by full replacement of natural sand by artificial sand as fine aggregate. Tests were carried out on cubes, cylinders and unreinforced beams to study the mechanical properties of concrete using artificial sand and compared with conventional concrete.

Keywords: Artificial sand, compressive strength, flexural strength, split tensile strength.

I. INTRODUCTION

Currently India has taken a major initiative on developing the infrastructures such as express highways, power projects and industrial structures etc. to meet the requirements of globalization, in the construction of buildings and other structures concrete plays the rightful role and a large quantum of concrete is being utilized. River sand, which is one of the constituents used in the production of conventional concrete, has become highly expensive and also scarce. In the backdrop of such a bleak atmosphere, there is large demand for alternative materials from industrial waste. Natural sand is excavated from river bed impacts on environment in many ways. Due to digging of the sand from river bed reduces the water head, so less percolation of rain water in ground, which result in lower ground water level. There is erosion of nearby land due to excess sand lifting as well as it destroys the flora & fauna in surrounding areas.

Due to limited supply of natural sand, cost is very high and its consistent supply cannot be guaranteed. Under these circumstances use of crushed fine aggregates becomes unavoidable. However, many people in India have doubts about quality of concrete and mortar with crushed sand as fine aggregates. As the supplies of suitable natural sand near the point consumption are becoming exhausted, the cost of this sand is increasing. In addition to this, the turbulence created by dredging sand near the estuaries could damage the fragile ecosystem along the coast. Thus a replacement material to the natural sand was sought, and the fines from crushing operations were identified as a possible substitute material by this research.

II. OBJECTIVES

The objective of this study is to investigate the mechanical properties of hardened concrete having artificial sand as fine aggregate

1. To assess conventional concrete.
2. To study the influence of artificial sand on the mechanical properties of concrete and compare the

result with that of concrete produced using selected river sand.

3. To clear doubts about quality and properties of concrete when artificial sand used as fine aggregates.

III. EXPERIMENTAL INVESTIGATION

3.1 Test materials and mix proportions

Portland pozzolana cement with ISI mark was used for tests on fresh and hardened concrete. The compressive strength was 40.75 MPa and 55.1 MPa at 7 and 28 days respectively. Local river sand and artificial sand with fineness modulus of 3.98 and 4.05 respectively were used. The maximum size in both the types of sand was 4.75 mm. The coarse aggregates with basaltic origin, maximum size 20 mm were from local stone crusher. Potable water, with pH of 7.1, was used. The designed mix M30 with proportion 1:1.37:2.75 (Cement: Fine aggregate: Coarse aggregate) for concrete with artificial sand and 1:1.30:2.81 for that with river sand on weight basis. The mix design was done as per IS 10262:2009. Water cement ratio of 0.435 kept constant for both the types of concrete and for all specimens.

3.2 Specimen Details

There were two series. Concrete with fine aggregate as river sand and concrete with fine aggregate as artificial sand and each series comprised of three beams. For each series six beams (150mm x 150mm x 1000mm), three cubes (150 mm x 150 mm x 150 mm) and three cylinders (150 mm diameter and 300 mm height) were cast as control specimens. Specimens were cured for 28 days.

3.3 Testing

Testing was carried out on 6 beams of both series for flexure. For flexural strength beams were simply supported on constant effective span of 900 mm under two point concentrated symmetrical loads for both series. All the beams were having constant overall span and width of 1000 mm and 150 mm respectively.



Figure 1: Flexural test setup

The beams were kept on universal testing machine. The beams were tested under gradually applied two point loading on Universal Testing machine (UTM) as shown in Fig. 1 for flexural strength. Ultimate load and modes of

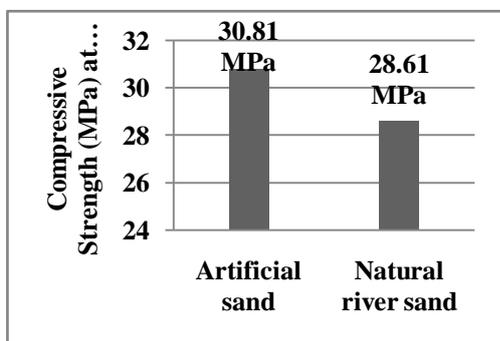
failure of beam were noted. Compressive strength and Split tensile strength are carried out on cubes and cylinders respectively, tested under compression testing machine.

IV. DISSCUSION AND RESULT

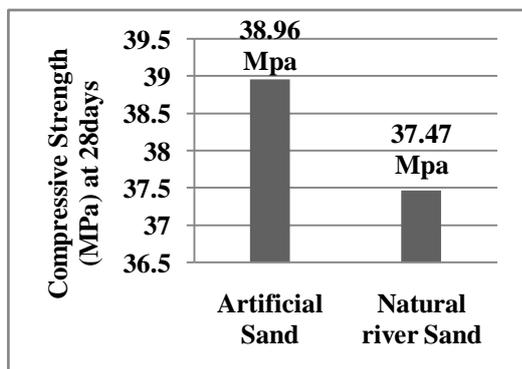
1.1 Compressive Strength

The calcium-silicate-hydrate (C-S-H) gel is the most important cementing component of concrete. It is responsible for the engineering properties of concrete including setting, hardening and strength development. The C-S-H gel micro-fibers penetrate in micro-pores of aggregate, it penetrates easily in crushed sand due to rough surface than river sand particles which results more strength.

From the result table, it is observed that, a compressive strength of cement concrete increases after replacing a fine aggregate by artificial / crushed sand at 7 and 28 days. During the crushing process the manufactured sand have irregular shapes and more fine particles which are filling gap between coarse particle and makes concrete solid also contributing to improved strength compared to river sand control mix.



Graph No.1: Compressive Strength (MPa) at 7 days

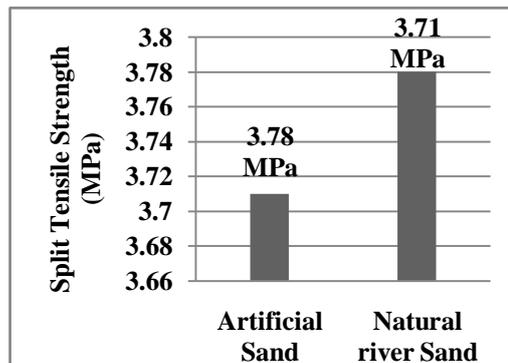


Graph No.2: Compressive Strength (MPa) at 28 days

1.2 Split tensile Strength

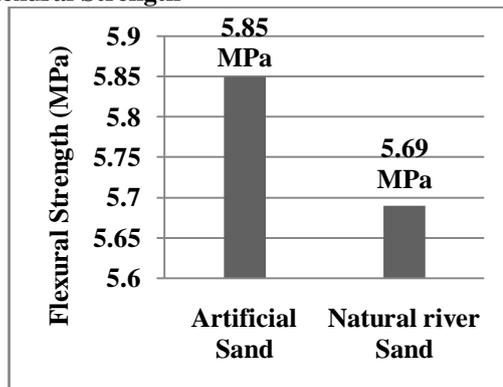
Plain concrete possesses a very low tensile strength, limited ductility and little resistance to cracking. Internal micro cracks are inherently present in the concrete and its poor tensile strength is due to propagation of such microcracks, eventually leading to brittle fracture of concrete.

From the experimental results it is seen that, the indirect split tensile strength of plain cement concrete with natural sand as fine aggregate is marginally higher than cement concrete with artificial sand as fine aggregate.



Graph No.3: Split Tensile Strength (MPa)

4.3 Flexural Strength



Graph No.4: Flexural Strength (MPa)

A good bond is essential for improving strength of concrete composite. The interfacial bond increased by larger area of contact, improving the frictional properties. The artificial sand contains more micro fines as compared to river sand which provides larger area of contact also particle shape of artificial sand is cubical, angular which helps to improve frictional properties that increases flexural strength of concrete.

Results have demonstrated that, the flexural strength of plain cement concrete with artificial / crushed sand as fine aggregate increased than cement concrete with natural river sand as fine aggregate. There was no crushing failure. All specimens are failed in pure bending zone, no shear cracks were seen.

V. CONCLUSIONS

1. A compressive strength of concrete with natural sand increased by 7.72% after fully replacing by artificial sand at 7 days and 3.98% at 28 days.
2. The effect on compressive strength of concrete by replacement of natural sand with artificial sand as fine aggregate is more prominent at seven days than that on 28 days.
3. The indirect split tensile strength of concrete with river sand as fine aggregate found marginally higher than concrete with artificial sand as fine aggregate, it is 3.78MPa and 3.71MPa respectively. Split tensile strength for all specimens was more than 10% of compressive strength.
4. The flexural strength of concrete with artificial sand as fine aggregate was recorded 2.81% more than concrete

with natural river sand as fine aggregate. All specimens are failed in pure bending zone of span, no shear failure were recorded.

REFERENCES

1. Gordana Toplicic, Curcic, Zoran Grdic, Iva Despotovic, Nenad Ristic, "Influence of crushed stone aggregate type on concrete consistency". *Architecture and Civil Engineering Vol. 8, No. 1, (2010), pp.99 – 109.*
2. Jafar Bolouri Bazaz, Mahamood Khayati, Navid Akrami, "Performance of concrete produced with crushed bricks as the coarse and fine aggregate". *The Geological Society of London 2006, IAEG2006, Paper number 616.*
3. B. Menadi a, S. Kenai a, J. Khatib b, A. Al t-Mokhtar, "Strength and durability of concrete incorporating crushed limestone sand". *Construction and Building Materials published by Science direct, , Vol. 23 (2009), pp. 625–633.*
4. H. Donza, O. Cabrera, E.F. Irassar, "High-strength concrete with different fine aggregate". *Cement and concrete research published by Pergamon, Vol. 32(2002), pp. 1755–1761.*
5. L.A. Balogun, D. Adepegba, "Effect of varying sand content in laterized concrete". *The international journal of cement composites and light weight concrete, Vol. 4, No.4 , pp.235-240.*
6. Tahir Celik, Khaled Marar, "Effect of crushed stone dust on some properties of concrete". *Cement and concrete research by Pergamon, Vol. 26, No. 7, pp.1121-1130 (1996).*
7. Kou Shi-Kong, Poon Chi-Sun, "Properties of concrete prepared with crushed fine stone, Furnace bottom ash and fine recycled aggregate as fine aggregates". *Construction and Building Materials published by Science direct, Vol. 23 (2009), pp. 2877-2886.*
8. Gurpreet Singh, Rafat Siddique, "Effect of waste foundry sand as partial replacement of sand on the strength, ultra sonic pulse velocity and permeability of concrete". *Construction and Building Materials published by Science direct, Vol. 26 (2012), pp. 416-422.*
9. IS 383. "Specifications for coarse and fine aggregates from natural sources for concrete", *Bureau of Indian Standards. New Delhi, 1970.*
10. IS 10262. "Recommended guidelines for concrete mix design", *Bureau of Indian Standards. New Delhi, 1982.*
11. IS 516. "Methods of tests for strength of concrete", *Bureau of Indian Standards. New Delhi, 1959.*
12. IS 456. "Code of practice for plain and Indian Standards". *Bureau of Indian Standards New Delhi, 2000.*

Table No.1: Compression test on at 7 days

Sr. No.	Type of Fine Aggregate	Compressive Load (N)	Compressive Strength (MPa)	Avg. Compressive Strength (MPa)
1.	Artificial sand	655×10^3	29.11	30.81
		749×10^3	33.29	
		676×10^3	30.04	
2.	Natural river sand	639×10^3	28.40	28.61
		575×10^3	30.00	
		617×10^3	27.42	

Table No.2: Compression test on cubes with river sand as fine aggregate at 28 days

Sr. No.	Type of Fine Aggregate	Compressive Load (N)	Compressive Strength (MPa)	Avg. Compressive Strength (MPa)
1.	Natural river sand	801×10^3	35.61	37.47
		897×10^3	39.87	
		831×10^3	36.93	
2.	Artificial sand	830×10^3	36.89	38.96
		895×10^3	39.78	
		905×10^3	40.22	

Table No.3: Split tensile test on cylinders at 28 Days

Sr. No.	Type of Fine Aggregate	Load (N)	Split Tensile Strength (MPa)	Avg. Split Tensile Strength (MPa)
1.	Artificial sand	228×10^3	3.23	3.71
		275×10^3	3.89	
		284×10^3	4.02	
2.	Natural river sand	247×10^3	3.50	3.78
		272×10^3	3.85	
		282×10^3	3.99	

Table No.4: Flexural strength test on beams at 28 days

Sr. No.	Type of Fine Aggregate	Load (N)	Flexural strength (MPa)	Avg. Flexural strength (MPa)
1.	Artificial sand	26.01×10^3	6.94	5.85
		20.05×10^3	5.35	
		19.80×10^3	5.28	
2.	Natural river sand	21.15×10^3	5.64	5.69
		20.90×10^3	5.58	
		21.95×10^3	5.85	

Structural complex configuration plate mathematical modeling and optimization

Sh. A. Nazirov, S. H. Yakubov

Tashkent University of information technologies, Karshi State University

Abstract: This paper examines the optimization problem of lamellar complex configuration structures. The results of a complex configuration plate structures' weight optimization calculation.

Keywords: Mathematical modeling, optimization, plate design, a complex form, methods of solution, R-function, various methods, numerical experiments, software package.

I. Introduction

In the design of various engineering structures, namely the construction of facilities, aircraft, missile, ship, etc. - there are problems of complex configuration (not rectangular plate shape, with cut-outs, multiply, etc.) lamellar structural elements calculation and optimization. The mathematical complexity of the calculation of these arbitrary shape plate elements, especially their optimization, resulted in a significant research and publications backlog on these issues from the calculation and optimization of the "tra

II. Statement of the Problem

The problem of engineering design optimizing will be put as mathematical programming problem: it is necessary to determine the vector $X(x_1, x_2, \dots, x_n)$ optimized parameters x_i ($i = 1, \bar{n}$), giving the objective function $F(x)$ extreme (for definiteness, we take min), keeping restrictions on the parameters $a_i \leq x_i \leq b_i$ ($i = 1, \bar{n}$) and functional limitations $f_j(x) \leq 0$ ($j = 1, \bar{m}$). This problem can be written

$$\begin{cases} F(X) \rightarrow \min, \\ f_j(X) \leq 0 \quad (j = 1, \bar{m}), \\ a_i \leq x_i \leq b_i \quad (i = 1, \bar{n}). \end{cases} \quad (1)$$

We shall consider equation (1) in details. The most commonly accepted parameters to be optimized in the structural elements are geometric (plate thickness h , curvature radius R_1 , external and internal edges, cutouts, etc.) and physical (elastic modulus E , etc.). The lower a_i and upper b_i meanings of x_i limits of parameters are defined on the basis of design and technological, operational, etc. requirements. For the objective function $F(X)$ the most commonly accepted parameters are: weight, materials consumption, the cost of construction.

The main functional constraints $f_j(X) \leq 0$ ($j = 1, \bar{m}$) for engineering structures, subjected to various external influences, optimization, are the following.

1. Stress state restrictions :

$$\max \sigma^{(\psi)}_{\text{KB}}(X) \leq [\sigma]^{(\psi)} \quad (\psi = 1, n). \quad (2)$$

Here ψ is the number of variants of the design impact; $\max \sigma^{(\psi)}_{\text{eq}}(X)$ - the maximum equivalent structural stress, defined according to the accepted hypothesis, or theory of strength, with ψ - m version of the impact, $[\sigma]^{(\psi)}$ - the allowable stress for the material of construction in the - m option exposure.

To the canonical form (1) restrictions (2) come as follows:

$$F_1(X) = \max \sigma^{(\psi)}_{\text{eq}}(X) - [\sigma]^{(\psi)} \leq 0.$$

2. Deformed state restrictions :

$$\max |u^{(\psi)}(X)| \leq [u]^{(\psi)},$$

where $\max |u^{(\psi)}(X)|$ - the maximum the surface structure displacement with m -option impact, $[u]^{(\psi)}$ - allowable surface structure displacement.

In the canonical form:

$$F_2(X) = \max |u^{(\psi)}(X)| - [u]^{(\psi)} \leq 0.$$

3. Stability conditions:

$$P\psi \leq P_{cr}$$

where $P\psi$ - compressive force with ψ -effect, P_{cr} - the critical force on the structure.

In the canonical form:

$$F_3(X) = P\psi - P_c \leq 0.$$

4. Restrictions on the natural oscillations frequency. The variable (periodic) loads activity at a certain frequency demands to analyze natural frequency constraints :

$$\min \{ w_i^{(\psi)}(X) \} \geq [w] \psi \quad (i = 1, 2, \dots),$$

where the $\min \{ w_i^{(\psi)}(X) \}$ is the lowest natural ψ -x oscillation frequency, $[w] \psi$ -the lowest allowed natural ψ -x oscillation frequency, appointed as the calculated value of the compelled ψ vibrations.

In the canonical form:

$$F_4(X) = [w] \psi - \min \{ w_i^{(\psi)}(X) \} \leq 0$$

5. Mechanical vibrations amplitude constraints:

$$a_0^{(\psi)}(X, w_i) \leq [a_0(w_i) \psi],$$

where $a_0^{(\psi)}(X, w_i)$ is the maximum forced ψ x- oscillations amplitude with the w_i frequency; and $[a_0(w_i) \psi]$ is the permissible amplitude.

The above mentioned restrictions are most common in the structures ‘ optimization, but the certain structures optimization solving tasks may require additional structural, technological, operational and other constraints.

Problem (1) of complex configuration structures engineering optimization is non-linear programming problem, which has a number of specific features. First, the calculation of the objective function (weight, cost) needs much less time than to check the restrictions, which require the construction calculation direct task solving, Second, the global minimum will always be at a border or at their junction, otherwise we will have a stockpile of material, that can be removed without violating the conditions of strength, stiffness, stability, etc. Third, the form of the $\sigma(X)$, $u(X)$, $P_c(X)$ etc. functions are a priori unknown and can only be determined numerically. Thus, to solve the problem (1) we shall apply the algorithms described in [2-15] taking into account the above features. The algorithm has high convergence speed and reliability.

III. Calculation methods

Let us consider the methods for solving the direct problem of calculation.

It is known that the equations of equilibrium, fluctuations and the stability of anisotropic plates, according to the moments are [1]:

$$\frac{\partial^2 M_1}{\partial x^2} + 2 \frac{\partial^2 M_{12}}{\partial x \partial y} + \frac{\partial^2 M_2}{\partial y^2} = q_1(x, y), \tag{3}$$

$$\frac{\partial^2 M_1}{\partial x^2} + 2 \frac{\partial^2 M_{12}}{\partial x \partial y} + \frac{\partial^2 M_2}{\partial y^2} + h(\sigma_x \frac{\partial^2 W}{\partial x^2} + \sigma_y \frac{\partial^2 W}{\partial y^2} + 2\sigma_{xy} \frac{\partial^2 W}{\partial x \partial y}) = 0 \tag{4},$$

$$\frac{\partial^2 M_1}{\partial x^2} + 2 \frac{\partial^2 M_{12}}{\partial x \partial y} + \frac{\partial^2 M_2}{\partial y^2} + m \frac{\partial^2 W}{\partial t^2} = q_2(x, y, t). \tag{5}$$

Here, W - plate deflection, M_1, M_{12}, M_2 - the bending and tensional moments, $m = \gamma h / g$, γ - the volume weight per unit, g - plate gravity acceleration, h - thickness.

Relations for the M_1, M_{12}, M_2 , when the plate is isotropic, orthotropic and anisotropic, are given in [1].

Substituting in (3), (4) the ratio of M_1, M_{12}, M_2 , when the plate is isotropic, orthotropic, or in other cases of anisotropy, it is possible to obtain the corresponding equations. These equations are given in many textbooks on the theory of elasticity [1-2].

Equations (3), (4) are supplied with the boundary conditions, and equation (5) –with both boundary and initial conditions.

Here are the types of encountered boundary conditions frequencies [1]:

a) rigidly clamped-edge

$$W|_{\Gamma=0} = 0, \quad \frac{\partial W}{\partial n}|_{\Gamma} = 0$$

b) free-simply supported edge

$$W|_{\Gamma=0}, \quad M_n|_{\Gamma} = (M_1 \cos^2 \alpha + M_{12} \cos \alpha \sin \alpha + M_2 \sin^2 \alpha)|_{\Gamma} = 0$$

where $\alpha = (n \wedge ox)$ and $\beta = (n \wedge oy)$ are the angles between the normals, relatively the axis Ox, Oy ;

c) the free edge

$$M_n|_{\Gamma=0}, \left(\theta_n + \frac{\partial}{\partial S} M_{nr}\right)|_{\Gamma} = 0$$

Where $\theta_n = M_1 \cos \alpha + M_2 \cos \beta$, $M_r = (M_2 - M_1) \cos \alpha \cos \beta + M_{12}(\cos^2 \alpha - \cos^2 \beta)$.

In addition, there are possible combinations of these boundary conditions, depending on the plates edges fixing method. The initial conditions for equations (5) have the form

$$W(x, y, t)|_{t=t_0} = W_0(x, y), \quad \dot{W}(x, y, t)|_{t=t_0} = \dot{W}_0(x, y).$$

The formation of the matrix to solve the above problems is carried out by V.L.Rvachev's R function [12] and Bubnov-Galerkin's [3-4] method's combination.

It should be noted that the direct application of the Bubnov-Galerkin method to solve equations (3), (4), (5) leads to computational difficulties. In this work further for the formation of resolving equations elements we shall use the method proposed in [12].

Here the application of the R – functions method is associated with the coordinate sequences construction, that will satisfy the boundary conditions without any approximations.

Coordinate sequences that satisfy the boundary conditions can be represented as an expansion

$$W = \sum_{i=1}^n T_i(t) B(\omega, \phi_i) = \sum_{i=1}^n T_i(t) W_i(x, y), \quad (6)$$

where $T_i(t)$ are unknown function of time, to be determined; $\{W_i(x,y)\}$ - a complete, linearly independent system of functions, which we will build, using V.L.Rvachev's R - functions method [12].

Note that in the case of static's in the representation (5) instead of $T_i(t)$ function the unknown coefficients C_i will occur. Substituting (6) to (3) - (5) and performing the usual procedure of the Bubnov-Galerkin method, we obtain the following equation:

$$AC=B, \quad (7)$$

$$A - \lambda B=0, \quad (8)$$

$$M\ddot{T} + AT = F, \quad (9)$$

$$T(t_0) = T_0, \quad \dot{T}(t_0) = \dot{T}_0,$$

where

$$b = \{b_{ij}\} = \frac{1}{S} \left\{ \iint_{\Omega} q_1 W_i d\Omega \right\}, \quad \{a_{ij}\} = \frac{1}{S} \left\{ \iint_{\Omega} f_{ij} d\Omega \right\}, \quad A =$$

$$B = \{b_{ij}\} = \frac{1}{S} \left\{ \iint_{\Omega} \varphi_{ij} d\Omega \right\}, \quad F = \{f_i\} = \frac{1}{S} \left\{ \iint_{\Omega} q_2 W_i d\Omega \right\},$$

$$M = \{m_{ij}\} = \frac{1}{S} \left\{ \iint_{\Omega} m W_i W_j d\Omega \right\}, \quad T(0) = \mu^{-1} T_1(t_0), \quad \dot{T}(0) = \mu^{-1} T_2(t_0)$$

$$T_1(t_0) = \frac{1}{S} \left\{ \iint_{\Omega} W_0 W_i d\Omega \right\}, \quad T_2(t_0) = \frac{1}{S} \left\{ \iint_{\Omega} \dot{W}_0 W_i d\Omega \right\},$$

$$f_{ij} = \left(\frac{\partial^2 W_i}{\partial x^2} + \nu \frac{\partial^2 W_i}{\partial y^2} \right) \frac{\partial^2 W_j}{\partial x^2} + \left(\frac{\partial^2 W_i}{\partial y^2} + \nu \frac{\partial^2 W_i}{\partial x^2} \right) \frac{\partial^2 W_j}{\partial y^2} + \frac{\partial^2 W_i}{\partial x \partial y} \cdot \frac{\partial^2 W_j}{\partial x \partial y}$$

$$\varphi_{ij} = \left(\sigma_x \frac{\partial^2 W_j}{\partial x^2} + \sigma_{xy} \frac{\partial^2 W_j}{\partial x \partial y} + \sigma_y \frac{\partial^2 W_j}{\partial y^2} \right) W_i,$$

To solve the system of equations (7) Gaussian elimination or the method of least squares and other methods depending on the properties of the matrix are applied. To determine the critical load the QL – method is applied.

We find the solution of equation (8) under condition (9) with the help of a variety of numerical methods: for example, by the central difference method or the Newmark method, or the method of quadrature sums or, others [4].

It should be noted that in the formation of the matrix, computation of the coordinate functions and their the n-th order derivatives' values, is carried out by the card operations [5, 12]. Here the integrals are computed by the n-point Gauss formula [4].

The above-described numerical algorithm allows to optimize the plate-like structures of both constant and variable thickness.

Thus, the computational algorithm of plate structures optimization consists of the following steps:

1. Objective function formation.
2. Functional limitations formation.
3. The parameters restrictions formation.
4. The direct calculation.
5. Strength, stiffness, stability and other conditions checking.

In its turn, the direct phase calculation consists of:

- Constructing a sequence of coordinate functions, satisfying the boundary conditions of the problem;
- The solving equation matrix elements formation ;
- The equation calculation.

It should be noted that the resolving equations can be algebraic or differential, depending on the problems considered in the static or dynamic formulation.

As mentioned above, the problem of engineering design optimizing will be put as a problem of mathematical programming.

Starting from the equation (1), we consider the optimization of weight plates, where $F(X)$ is the weight of a plate of isotropic material under the action of the external load q . Functional limitations, taking into account in the engineering designs optimization, as well as a numerical optimization algorithm of complex configuration lamellar structures are described in detail in [4-9,11,14,15]. As the optimized option we take the plate thickness, constant in the plate range.

IV. Experimental calculations.

Task 1. The tightly clamped round the whole contour plate under uniform external pressure $q = 10$ kg. weight optimization. The radius of the plate is $R = 100$ cm, elastic modulus is $E = 2^{10} \cdot 6 \text{ kg/cm}^2$, Poisson's ratio is $\nu = 0.3$, permitted deflection is $[W] = 1$ cm and equivalent stress is $[\sigma_{\hat{y}\hat{e}\hat{a}}] = 2550 \hat{e}\hat{a}/\hat{n}\hat{i}^2$, the gravity (specific weight) is $\gamma = 7 \cdot 8 \text{ g/cm}^3$, $G(h) = \pi R^2 \gamma h \rightarrow \min$, $W_{\max} \leq [W]$, $\sigma_{\max}^{\text{экв}} \leq [\sigma_{\text{экв}}]$, $1 \text{ cm} \leq h \leq 10 \text{ cm}$.

Optimization was carried out up to $\varepsilon = 0.01$. We obtained the following results:

$$G_{\min} = 1141.19 \text{ kg}, \quad h = 4.6571 \text{ cm},$$

$$W_{\max} = 0.8206 \text{ cm}, \quad \sigma_{\max}^{\text{экв}} = 2549.99 \text{ kg/cm}^2.$$

This problem has an exact solution:

$$W_{\max} = \frac{qR^4 12(1-\nu^2)}{64Eh^3}.$$

With the calculated value of h , we have :

$$W_{\max} = 0.8446 \text{ cm}.$$

The accuracy of the obtained approximate solution is satisfactory.

Task 2. Optimization of the entire ring rigidly clamped at both the contours plate under uniform external pressure intensity $q = 10 \text{ kg/cm}^2$.

The outer radius of the plate $R = 100$ cm, inner - $r = 50$ cm. The other parameters are the same as in Task 1:

$$G(h) = \pi \gamma h (R^2 - r^2) \rightarrow \min$$

The results are :

$$G_{\min} = 382.95 \hat{e}\hat{a}, \quad h = 2.0837 \hat{n}\hat{i}, \quad \sigma_{\max}^{\hat{y}\hat{e}\hat{a}} = 2549.82 \hat{e}\hat{a}/\hat{n}\hat{i}^2$$

Task 3. Optimization round the whole ring (Fig. 1). All parameters are the same as in Task 2. We obtained the following results:

$$G_{\min} = 134.48 \text{ kg}, \quad h = 1.126 \text{ cm}, \quad \sigma_{\max}^{\text{экв}} = 2549.99 \text{ kg/cm}^2$$

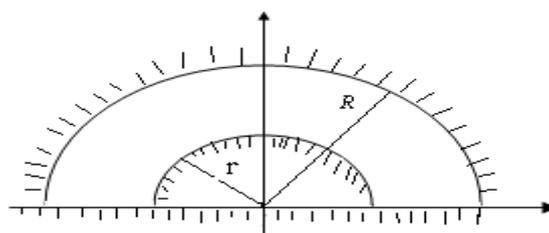


Fig. 1

Task 4. A square with round neck weight optimization (Fig. 2).

Initial data for the square with a round neck weight optimization calculation:

$$G(h) = (ab - \pi r^2) \gamma h, \quad a = 200 \text{ cm}, \quad b = 200 \text{ cm}, \quad r = 50 \text{ cm}, \quad q = 10 \text{ kg} / \text{cm}^2,$$

$$W_{\max} = [W], \quad \sigma_{\max}^{\text{opt}} \leq [\sigma_{\text{opt}}], \quad 1 \text{ cm} \leq h \leq 10 \text{ cm}.$$

The results of the calculation: $G_{\min} = 296.07 \text{ kg}$, $h = 3.9767 \text{ cm}$, $\sigma_{\max}^{\text{opt}} = 2546.07 \text{ kg} / \text{cm}^2$.

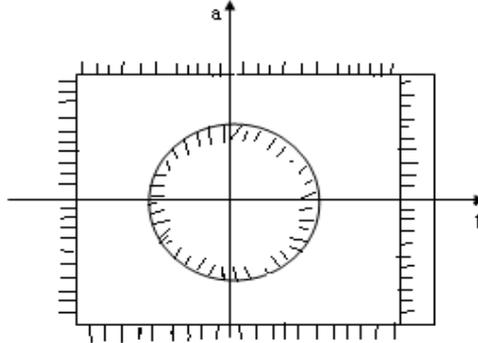


Fig.2

Task 5. Weight optimization of the figure shown in Fig. 3 Baseline data:

$$G(h) = (ab - 3\pi r^2) \gamma h, \quad a = 200 \text{ mm}, \quad b = 200 \text{ mm}, \quad r = 20 \text{ mm},$$

$$W_{\max} \leq [W], \quad \sigma_{\max}^{\text{opt}} \leq [\sigma_{\text{opt}}], \quad 1 \text{ mm} \leq h \leq 10 \text{ mm}.$$

We obtained the following results:

$$G_{\min} = 725.96 \text{ kg}, \quad h = 2.5689 \text{ cm}, \quad W_{\max} = 0.1262 \text{ cm}, \quad \sigma_{\max}^{\text{opt}} = 2549.9998 \text{ kg} / \text{cm}^2.$$

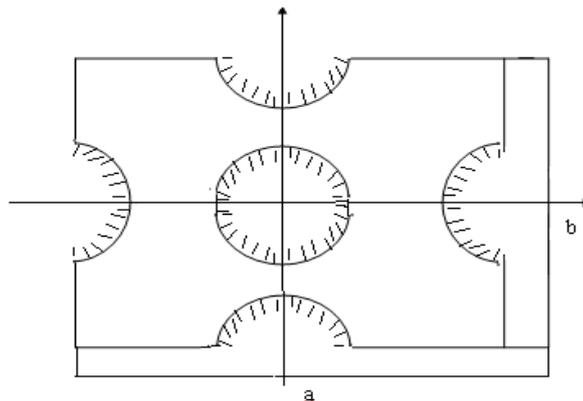


Fig. 3

Tasks 1 and 2 have the exact solutions and are given only for the algorithm [5] performance monitoring possibility. According to the solved problems, the main limitation (with the taken values $[\sigma_{\text{opt}}]$, $[\sigma] u [W]$) is the strength limitation, and the algorithm provided a high degree approximation to the boundary. A variety of forms of plates indicates wide opportunity of applying the algorithm to solve optimization problems of plates of complex configuration [4, 15].

V. Conclusion

Thus, the proposed technique allows to optimize on the weight the calculation experiments of complex configuration design plates.

VI. References

- [1] Ambartsumyan S.A. Theory of anisotropic plates: strength, stability and fluctuations. // 2nd ed. revised. and add. - Moscow: Nauka, 1987. - 360 p.
- [2] Vlasov V.Z. The general theory of shells. -M. Fizmatgiz, 1962.-T. A.
- [3] Kabulov V.K. Algorithmic in continuum mechanics Avg.-Fan, Tashkent, 1980.-304 p.
- [4] Kabulov V.K., Nazirov SH.A., Yakubov S.H. Algorithm of solving optimization problems. - Fan, Tashkent, 2008. - 204 p.
- [5] Nazirov SH.A., Piskorskii L.F. A numerical algorithm for optimization of plate structures with a complex configuration // Questions of High and Appl. Mathematics.": Science works coll., Tashkent, IR AS RU, 1995. - Issue 100. - C. 49-56.

- [6] Nazirov SH.A., Piskorskii L.F. Software for calculation and optimization of engineering structures of complex configuration // Mathematical modeling and computational experiment Conf. report. Tashkent, 1994 - C. 215.
- [7] Nazirov SH.A., Piskorskii L.F. Software for calculation and optimization of plate structures with complex shape // Algorithms. Science works coll. - Tashkent, IR AS RU, 1995. - Vyp.80. - C. 41-54.
- [8] Nazirov SH.A., Piskorskii L.F, Bobokulov SH.O, Ngo Ngoc Hung. Optimization of elastic plates of complex configuration // Issues. Comput. and Appl. Mathematics. Science research, Tashkent, IR AS RU, 1995. - Vyp.99. - C. 41-45.
- [9] Nazirov Sh.A, Piskorskii L.F, Bobokulov SH.O. Algorithmic optimization problems in the theory of elasticity // "Problems of Informatics and Energy" magazine. Tashkent, 1997. - № 3.
- [10] .Nazirov SH.A., Yakubov S.H. The algorithmic system that automates the process of optimization for the design of engineering structures and buildings // The Rep. Uz. State Patent Office. Certificate, DGU 01422. 13.11. In 2007.
- [11]. Nazirov S.A, Yakubov S.H. The development of algorithmic solutions of classes of optimization problems // Strategy and the development of science and technology in the twenty-first century. Proceedings of the science practical conference. Bukhara, 2009. V.2. - S. 91-94.
- [12] Rvachev V.L, Kurpa L.V. R - function in the plate theory - Kiev: Naukova Dumka, 1987. - 176 p.
- [13] .Samarsky A.A., Nikolaev E.S. Methods for net (grid) equations solving. - Moscow: Nauka, 1978. -592 p
- [14] akubov S.H. On the intelligent systems for the optimal design of engineering structures of complex configuration development // Proceedings of the conference. Karshi State University-Karshi: Nasaf, 2009. - S. 159-162.
- [15] .Yakubov S.H. The engineering structures and buildings design optimization system analysis // Complex systems' optimization problems The seventh Intern. Asian Summer School report. Tashkent, 2011. - P.154-163.

Performance - Detection of Bad Smells In Code for Refactoring Methods

Karnam Sreenu¹, D. B. Jagannadha Rao²

¹, (Assistant Professor, Department of Information Technology, Sreenidhi Institute of Science and Technology, India

², (Associate Professor, Department of Computer Applications, Institute of Science and Technology, India

ABSTRACT: Day by day the complexity levels of Software system increasing. Hence more effort is required for software organizations to develop new or rebuild existing system of high quality. Refactoring reduces the cost of software maintenance through changing the internal structure of the source-code to improve the overall design that helps the present and future developers to evolve and understand a system. This paper describes new refactoring methods and metrics along with the existing metrics to identify the characteristics of bad smells "Lazy Class" and "Temporary Field" through which the developer can be provided with significant guidance to locate bad smells. After identifying these bad smells, appropriate refactoring methods can remove them.

Keywords: Software Refactoring, Bad Smells, Software Metrics, Software Quality.

I. INTRODUCTION

The design of source-code has become an increasingly important part of the overall development of software. Refactoring changes the internal code structure of an Object-Oriented (O-O) system without affecting the overall behavior of the system to improve the quality of the design [1]. Refactoring is a process of making semantic-preserving transformations of code into a form that the software engineer finds easier to understand.

Refactoring or the restructuring of a software system without changing its behavior is necessary to remove quality defects that are introduced by quick and often unsystematic development.

Refactoring is starting to become an integrated part of other software development processes to improve the design, help make design changes, integrate new functionality, and help understand the underlying design concepts.

The process of refactoring has three distinct stages to its application: identify where to apply a refactoring, choose an appropriate refactoring as a solution and apply the refactoring. Current software tools and Fowler's description of refactorings only consider the final stage of applying refactoring methods automatically and manually. Knowing where an appropriate place and which refactorings to apply in a system is arguably quite difficult.

One particular motivation is to improve the design of a software system through locating problems in the design and using refactoring as a solution.

Fowler and Beck [1] defined bad smells that describe a design problem that have a number of related Refactorings that can change the structure of a system to help improve the design. However locating bad smells currently involves manually inspecting source-code, which quickly becomes unfeasible as the size of the system

Increases. Providing an automatic support for the detection of bad smells becomes quite appealing.

The motivation for this paper is to enhance the well established refactoring process of identifying where to apply refactorings in a system. The focus will be on automatically identifying bad smell design problems in Java source code. To achieve the goal a prototype tool is developed that applies a set of software metrics on Java systems and the results are interpreted to identify problems in the design (i.e. bad smells).

II. BACKGROUND

2.1 BAD SMELLS

Some of the bad smells from Fowler's book [1] are summarized below:

- **Duplicate Code:** The same code structure in two or more places is a good sign that the code needs to be refactored: if you need to make a change in one place, you will probably need to change the other one as well, but you might miss it.
- **Long Method:** Long methods should be decomposed for clarity and ease of maintenance.
- **Long Parameter List:** Long parameter lists are hard to understand. You don't need to pass in everything a method needs, just enough so it can find all it needs.
- **Shotgun Surgery:** If a type of program change requires lots of little code changes in various different classes, it may be hard to find all the right places that do need changing. May be the places that are affected should all be brought together into one class.
- **Feature Envy:** This is where a method on one class seems more interested in the attributes (usually data) of another class than in its own class, May be the method would be happier in the other class.
- **Large Class:** Classes that are trying to do too much often have large numbers of instance variables.
- **Data Class:** Classes that just have data fields, and access methods, but no real behaviour. If the data is public, make it private.
- **Lazy Class:** Classes that are not doing much useful work should be eliminated.
- **Temporary Field:** It can be confusing when some of the member variables in a class are only used occasionally.

2.2 REFACTORING METHODS

- **Push down method:** 'Behavior on a super class is relevant only for some of its subclasses'. The method is moved to those subclasses.
- **Pull up Method:** 'You have methods with identical results on subclasses'. In this case, the methods should be moved to the super class.

- **Pull up field:** ‘Two subclasses have the same field’. In this case, the field in question should be moved to the super class.
- **Move field:** ‘A field is, or will be, used by another class more than the class on which it is defined’.
- **Rename Method:** A method is renamed to make its purpose more obvious.
- **Rename Field:** A field is renamed to make its purpose more obvious.
- **Move Method:** ‘A method is, or will be, using or used by more features of another class than the class on which it is defined’.

2.3 REFACTORING PROCESS

Refactoring can be divided into a number of steps as shown below [2]:

1. Identify where the software needs to be refactored.
2. Determine which refactorings need to be applied to the identified places.
3. Guarantee that the applied refactoring preserves behavior.
4. Apply the refactoring.
5. Assess the effect of the refactoring on the quality characteristics of the software or the process.
6. Maintain the consistency between the refactored program code and other software artifacts.

III. PROPOSED WORK

This paper provides two new refactoring methods called Merge Class Refactoring and Replace Temp Refactoring, also describes two metrics called Number of Methods (NOM) and Instance Variable per Method in a Class (IVMC). These two refactoring methods are mainly used to reduce the lines of source code.

Also provides a bad smell description framework and bad smell interpretation framework to collect the information regarding bad smells. These frameworks mainly contain three parts.

Bad Smell Description Framework:

- **Bad Smell Name:** It is the description of the bad smell which is proposed by Fowler and Beck’s.
- **Characteristics of bad smell:** Identifying main characteristics from description of the above bad smell.
- **Identifying any design heuristics from the characteristics.**

Bad Smell Interpretation Framework:

- **Bad Smell Name:** It is the description of the bad smell which is proposed by Fowler and Beck’s.
- **Measurement Process:** Describe possible measurement techniques that when applied to Java source-code can help identify the design problem.
- **Interpretation Rules:** The interpretation indicates a set of rules on how the metrics can be used to identify possible candidates. We are using conventional metrics and new metrics to identify bad smells “Lazy Class” and “Temporary Field”.

3.1 REFACTORING MODEL:

Figure 1 describes the detail process about how the bad smells are identified in the source code and determining

which refactoring can be applied with the help of metrics values, then we can apply the appropriate refactoring method on the source code.

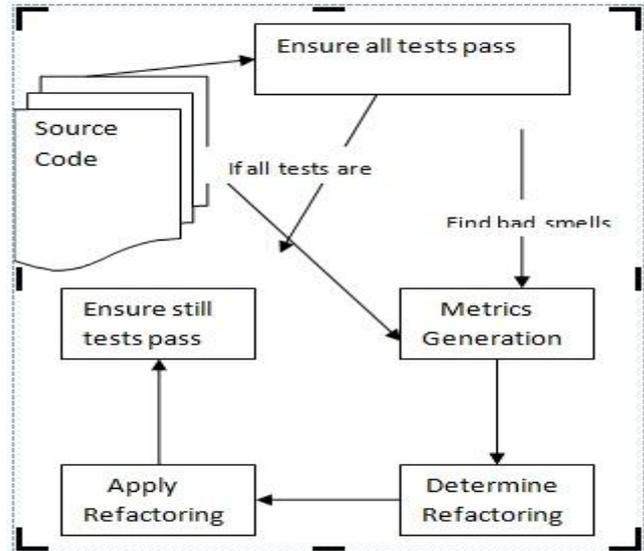


Figure 1: Describes the detail process

3.1.1. MERGE CLASS REFACTORING:

To apply this refactoring on the source code, first we have to identify the class to merge with the targeted class. To do this refactoring we are calculating some metric values to find out the lazy class which is not doing much work.

Lazy Class: To identify this bad smell the following are possible interpretation rules:

- if $NOM = 0$
- if $(LOC < LOCThreshold)$
- if $(DIT > 1)$

LOC: Lines of Code

DIT: Depth of Inheritance

If the above rules are true then we can directly apply the merge class refactoring. That is the class which satisfies the above conditions can be merged with the targeted class.

Example:

```

Class Person
{
    String name;
    int getTelNumber();
}
Class TelNumber
{
    int areacode;
    String number;
    int getTelNumber();
    int getAreaCode();
}
  
```

On the above two classes we can apply the metrics to find out the lazy class, in which the lazy class is class “Person”. So, it is merged with targeted class name called “Person”.

```

Class Person
{
    String name;
    int areacode;
    String number;
    int getTelNumber();
}
  
```

```
int getAreaCode();  
}
```

Conditions:

I. If some classes have the same number of methods (NOM) then we should calculate Lines of Code for (LOC) those classes to know or decide which class has to be merged with targeted class.

II. The Merge Class Refactoring Method can be directly applied on the source code by calculating the Number of Methods (NOM) metric when there is no inheritance mechanism in the source code.

IV. We should calculate the Depth of Inheritance (DIT) metric to apply the Merge Class Refactoring Method on the source code when the inheritance mechanism will present in the source code.

3.1.2. REPLACE TEMP REFACTORING

We have a temp that is assigned to once with a simple expression, and the temp is getting in the way of other refactoring's. Replace all references to that temp with the expression.

To apply this refactoring on the source code, first we have to identify the instance variable, that is not important in the source code. To do this the following interpretation rule is used.

if $IVMC \leq 2$

The above statement will be true when the instance variables which are declared in the code are not used more than two times. If the above rule is satisfied then we are ready apply the replace temp refactoring on the source code. By applying this refactoring on the source code we can remove the unused or unimportant instance variables so that the lines of the source code will be decreased.

Example1:

```
double random = ran.number();  
return (random <= 1);
```

In the above code "random" variable is treated as temporary variable, without its presence also the program is working correctly without changing its external behaviour. So, after applying the refactoring method the code will modified as following:

```
return (ran.number() <= 1);
```

Example2:

The following code snippet also shows how the replace temp refactoring will occur to remove the instance variables which are not used more than two times.

Before Refactoring

```
int a=10, b=20, c=0;  
c=a + b;
```

After First Refactoring

```
int a=10, b=20, c=0;  
c=10 + b;
```

After Second Refactoring

```
int a=10, b=20, c=0;  
c=10 + 20;
```

Conditions:

I. To apply this refactoring on the source code, first we have to identify the instance variable, that is not important in the source code.

II. The $IVMC \leq 2$ condition will be true when the instance variables which are declared in the code are not used more than two times.

V. IMPLEMENTATION

We have implemented the above mentioned metrics (NOM, LOC, DIT, and IVMC) in java [5] to find out the bad smells in the source code. If number of methods are equal in "Class A" and "Class B" then we have to calculate the Lines of Code for "Class A" and "Class B", based on these two values we have applied the appropriate Refactoring method. By using these implemented metric values we have applied the appropriate Refactoring method on the source code to remove the bad smell from the existing code or to improve the structure of the existing source code.

VI. CONCLUSION

The Merge Class Refactoring method and Replace Temp Refactoring method will be identified by providing metric values based on Number of Methods (NOM), Instance Variable per Method in Class (IVMC) and some existing metrics like Lines of Code (LOC), Depth of Inheritance (DIT) [3]. By identifying these metric values we can apply the above two refactoring methods directly on the source code to reduce the total number of lines of code (LOC) and to improve the structure of existing code.

REFERENCES

- [1] Fowler M., Refactoring: Improving the Design of Existing Code, 1st ed: Addison-Wesley, ISBN 0-201-48567-2, 1999.
- [2] T. Mens, T. Tourwe. "A Survey of Software Refactoring," IEEE Transactions on Software Engineering, Vol. 30, No.2, February 2004.
- [3] Shyam R. Chidamber and Chris F. Kemerer "A Metrics Suite for Object Oriented Design" IEEE Transactions on Software Engineering, vol. 20, no. 6, June 1994.
- [4] Rapu D., Ducasse S., Girba T., and Marinescu R., "Using history information to improve design flaws detection," presented at Eighth European Conference on Software Maintenance and Reengineering, Tampere, Finland, 2004, pp. 223-232.
- [5] Naughton Schildt, The Complete Reference Java2 Third Edition.

AUTHOR'S PROFILE



Mr. Karnam Sreenu has completed his M. Tech in Software Engineering from College of Engineering, Ananthapur. He is having around four years of experience. He is working as Assistant Professor in Sreenidhi Institute of Science and Technology, Ghatakeswar, Hyderabad. He is mainly interested in Software Engineering and algorithms.



Mr. D. B. Jagannadha Rao has completed his M. Tech in Computer Science from College of Engineering, Ananthapur. He is having around nine years of experience. He is working as Associate Professor in Sreenidhi Institute of Science and Technology, Ghatakeswar, Hyderabad. He is especially interested in Mobile Ad-hoc Networks, Software Engineering and its routing algorithms.

A Five Level Inverter for Grid Connected PV System Employing Fuzzy Controller

M. Pavan Kumar¹, A. Sri Hari Babu²

^{1,2} (Department of Electrical and Electronics Engineering, VIGNAN University, INDIA)

ABSTRACT: This paper presents a single-phase five-level photovoltaic (PV) inverter topology for grid-connected PV systems with a novel pulse width-modulated (PWM) control scheme. Two reference signals identical to each other with an offset equivalent to the amplitude of the triangular carrier signal were used to generate PWM signals for the switches. The inverter offers much less total harmonic distortion and can operate at near-unity power factor. This paper presented a single-phase multilevel inverter for PV application. It utilizes two reference signals and a carrier signal to generate PWM switching signals. The circuit topology, modulation law, and operational principle of the proposed inverter were analyzed in detail. A FUZZY control is implemented to optimize the performance of the inverter. MATLAB/SIMULINK results indicate that the THD of the Fuzzy Controller Circuit is much lesser than that of the THD of the PI Controller Circuit. Furthermore, both the grid voltage and the grid current are in phase at near-unity power factor.

Keywords: Grid connected Photovoltaic system, Maximum power point tracking system, Single phase five level inverter and fuzzy logic controller.

I. INTRODUCTION

The Demand for renewable energy has increased significantly over the years because of shortage of fossil fuels and greenhouse effect. The definition of renewable energy includes any type of energy generated from natural resources that is infinite or constantly renewed. Examples of renewable energy include solar, wind, and hydropower. Renewable energy, due to its free availability and its clean and renewable character, ranks as the most promising renewable energy resources like Solar energy, Wind energy that could play a key role in solving the worldwide energy crisis. Among various types of renewable energy sources, solar energy and wind energy have become very popular and demanding due to advancement in power electronics techniques. Photovoltaic (PV) sources are used today in many applications as they have the advantages of effective maintenance and pollution free. Solar electric energy demand has grown consistently by 20% to 25% per annum over the past 20 years, which is mainly due to its decreasing costs and prices. This decline has been driven by the following factors.

- 1) An increasing efficiency of solar cells,
- 2) Manufacturing technology improvements,
- 3) Economies of scale.

PV inverter, which is the heart of a PV system, is used to convert dc power obtained from PV modules into ac power to be fed into the grid. Improving the output waveform of the inverter reduces its respective harmonic content and, hence, the size of the filter used and the level of Electromagnetic Interference (EMI) generated by switching operation of the inverter. In recent years, multilevel inverters have become more attractive for researchers and manufacturers due to their advantages over conventional three level PWM Inverters. They offer improved output waveforms, smaller filter size and lower EMI, lower Total Harmonic Distortion (THD). The three common topologies for multilevel inverters are as follows:

- 1) Diode clamped (neutral clamped),
- 2) Capacitor clamped (flying capacitors),
- 3) Cascaded H-bridge inverter.

In addition, several modulation and control strategies have been developed or adopted for multilevel inverters, including the following multilevel sinusoidal (PWM), multilevel selective harmonic elimination, & Space Vector modulation. A typical single phase three-level inverter adopts full-bridge configuration by using approximate sinusoidal modulation technique as the power circuits. The output voltage then has the following three values: zero, positive (+Vdc), and negative(-V dc) supply dc voltage (assuming that Vdc is the supply voltage). The harmonic components of the output voltage are determined by the carrier frequency and switching functions. Therefore, their Harmonic reduction is limited to a certain degree. To overcome this limitation, this paper presents a five-level PWM inverter whose output voltage can be represented in the following five levels: zero, +1/2Vdc, Vdc, -1/2V dc, and -V dc. As the number of output levels increases, the harmonic content can be reduced. This inverter topology uses two reference signals, instead of one reference signal, to generate PWM signals for the switches. Both the reference signals Vref1 and Vref2 are identical to each other, except for an offset value equivalent to the amplitude of the carrier signal Vcarrier, as shown in Fig.1.

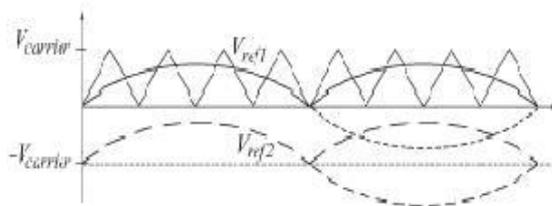


Fig.1 Carrier and reference signals

Because the inverter is used in a PV system, a Fuzzy control scheme is employed to keep the output current sinusoidal and to have high dynamic performance under rapidly changing atmospheric conditions and to maintain the power factor at near unity. Simulation results are presented to validate the proposed inverter configuration.

II. FIVE LEVEL INVERTER TOPOLOGY AND PWM LAW

The proposed single phase five level inverter topology is shown in Fig.2. The inverter adopts a full-bridge configuration with an auxiliary circuit. PV arrays are connected to the inverter via a dc–dc Boost converter.

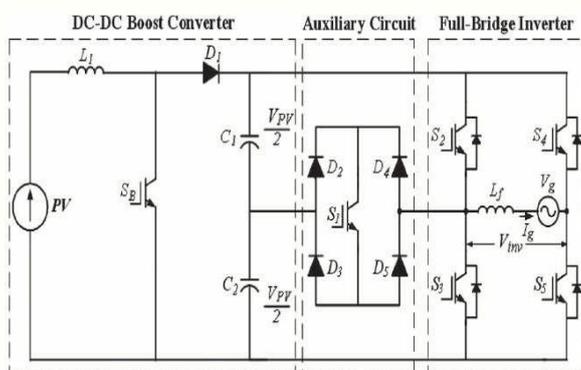


Fig.2 Single phase five level inverter topology

Because the proposed inverter is used in a grid-connected PV system, utility grid is used instead of load. The dc–dc boost converter is to be used to step up inverter output voltage V_{in} to be more than 2 of grid voltage V_g to ensure power flow from the PV arrays into the grid.

A filtering inductance L_f is used to filter the current injected into the grid. The injected current must be sinusoidal with low harmonic distortion. In order to generate sinusoidal current, sinusoidal PWM is used because it is one of the most effective methods. Sinusoidal PWM is obtained by comparing a high frequency carrier with a low frequency sinusoid, which is the modulating or reference signal. The carrier has a constant period; therefore, the switches have constant switching frequency. The switching instant is determined from the crossing of the carrier and the modulating signal.

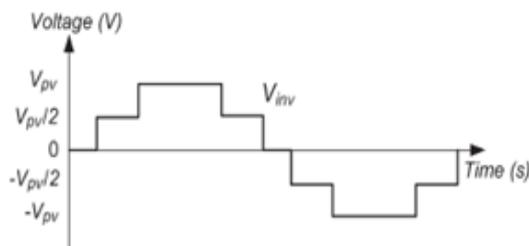


Fig.3 Ideal five-level inverter output voltage V_{inv} .

III. OPERATIONAL PRINCIPLE OF THE PROPOSED INVERTER

Because PV arrays are used as input voltage sources, the voltage produced by the arrays is known as V_{arrays} . V_{arrays} boosted by a dc–dc boost Converter to exceed $2V_g$. The voltage across the dc-bus capacitors is known as V_{pv} . The operational principle of the proposed inverter is to generate five level output voltage, i.e., 0 , $+V_{pv}/2$, $+V_{pv}$, $-V_{pv}/2$, and $-V_{pv}$. Proper switching control of the auxiliary circuit can generate half level of PV. Supply voltage, i.e., $+V_{pv}/2$, $+V_{pv}$, $-V_{pv}/2$. Two reference signals V_{ref1} and V_{ref2} will take turns to be compared with the carrier signal at a time. If V_{ref1} exceeds the peak amplitude of the carrier signal $V_{carrier}$, V_{ref2} will be compared with the carrier signal until it reaches zero. At this point onward, V_{ref1} takes over the comparison process until it exceeds $V_{carrier}$. This will lead to a

switching pattern, as shown in Fig. 4. Switches S1–S3 will be switching at the rate of the carrier signal frequency, whereas S4 and S5 will operate at a frequency equivalent to the fundamental frequency. Table I illustrates the level of V_{inv} during S1–S5 switch on and off.

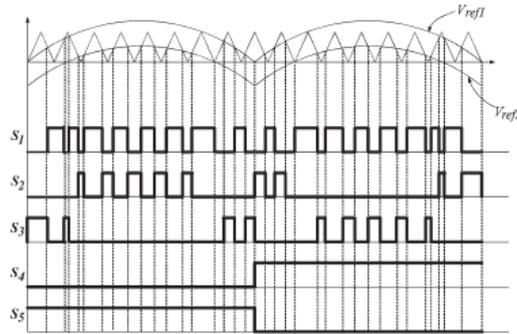


Fig.4 switching pattern for the single-phase five-level inverter

IV. CONTROL SYSTEM ALGORITHM AND IMPLEMENTATION

The feedback controller used in this application utilizes the FUZZY algorithm. As shown in Fig., the current injected into the grid, also known as grid current I_g , is sensed and fed back to a comparator which compares it with the reference current I_{ref} . I_{ref} is obtained by sensing the grid voltage and converting it to reference current and multiplying it with constant m . This is to ensure that I_g is in phase with grid voltage V_g and always at near-unity power factor. One of the problems in the PV generation systems is the amount of the electric power generated by solar arrays always changing with weather conditions, i.e., the intensity of the solar radiation.

TABLE I: Inverter Output Voltage during S1-S5 Switch ON and OFF

S1	S2	S3	S4	S5	V_{inv}
ON	OFF	OFF	OFF	ON	$+V_{pv}/2$
OFF	ON	OFF	OFF	ON	$+V_{pv}$
OFF	OFF	OFF	ON	ON	0
ON	OFF	OFF	ON	OFF	$-V_{pv}/2$
OFF	OFF	ON	ON	OFF	V_{pv}

A maximum power point tracking (MPPT) method or algorithm, which has quick-response characteristics and is able to make good use of the electric power generated in any weather, is needed to solve the aforementioned problem. Constant m is derived from the MPPT algorithm. The perturb and observe algorithm is used to extract maximum power from PV arrays and deliver it to the inverter. The instantaneous current error is fed to a FUZZY controller. The integral term in the FUZZY controller improves the tracking by reducing the instantaneous error between the reference and the actual current. The resulting error signal u which forms V_{ref1} and V_{ref2} is compared with a triangular carrier signal and intersections are sought to produce PWM signals for the inverter switches.

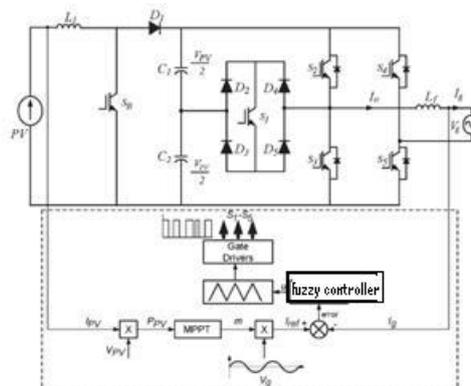


Fig.5 Five level inverter with FUZZY control.

The Trapezoidal sum approximation is used to transform the integral term into the discrete time domain because it is the most straightforward technique. The proportional term is directly used with out approximation. The Simulation result for five level inverter for grid connected PV system

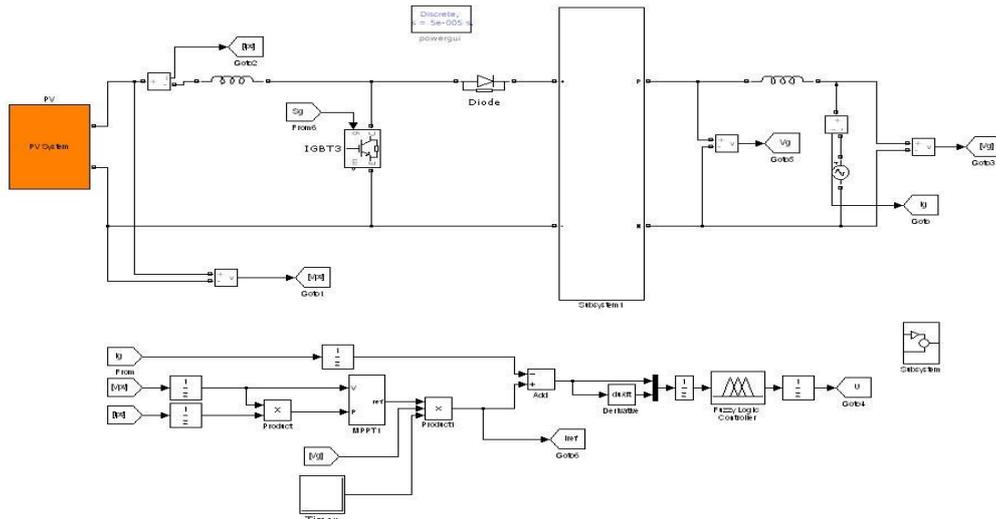


Fig.7 Simulink diagram for five level inverter

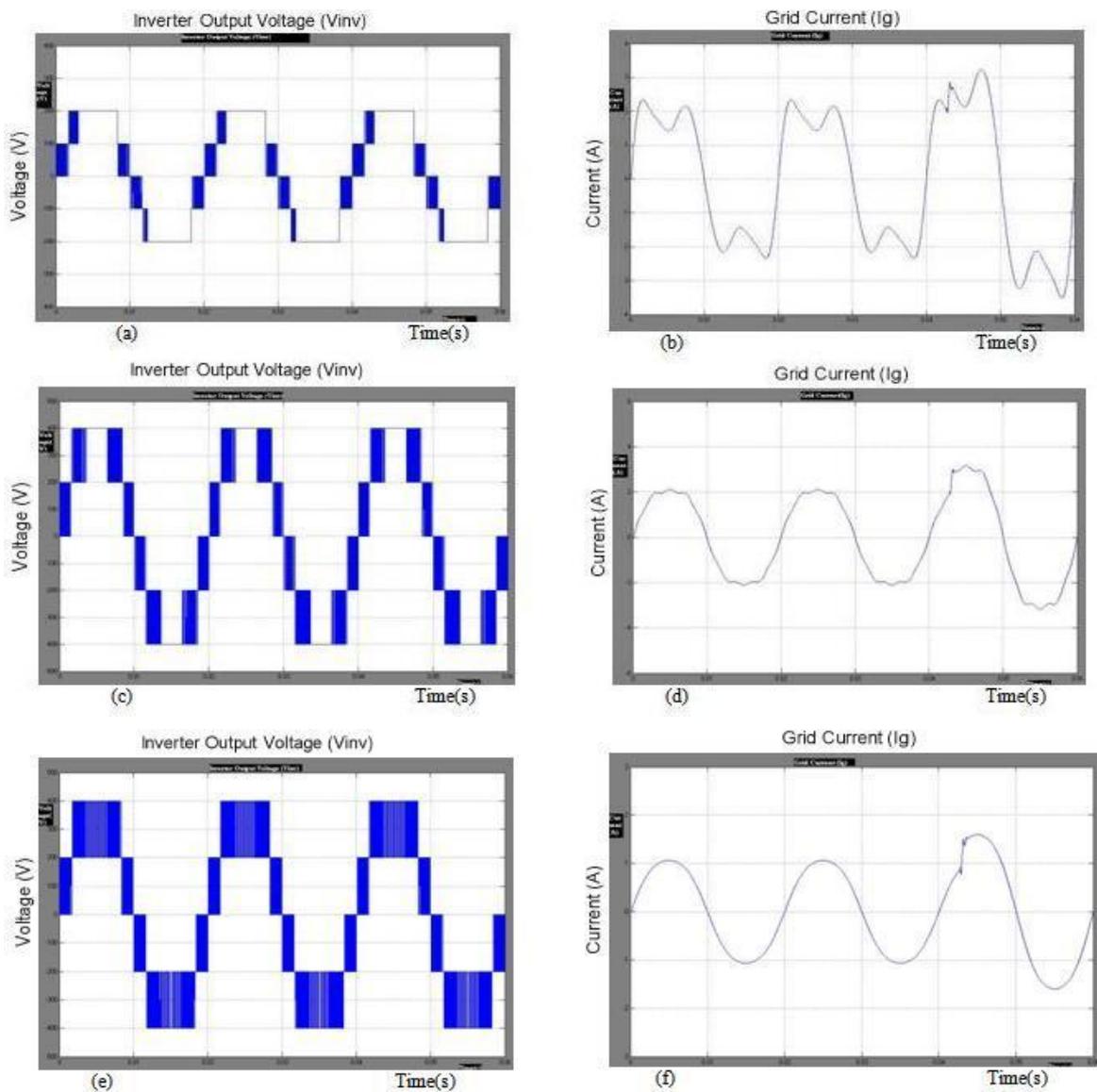


Fig.7 Inverter output voltage (V_{inv}) and grid current (I_g) for different values of M .

(a) V_{inv} for $M < 0.5$. (b) I_g for $M < 0.5$. (c) V_{inv} for $M > 1.0$. (d) I_g for $M > 1.0$. (e) V_{inv} for $0.5 \leq M \leq 1.0$. (f) I_g for $0.5 \leq M \leq 1.0$.

Table II: THD Analysis Table

Modulation Index(M)	Using Fuzzy Controller
$M < 0.5$	39.94%
$M > 1.0$	9.39%
$0.5 \leq M \leq 1.0$	4.65%

Table III: PV Multilevel Inverter Specifications and Controller Parameter

S1-S5 & SB	IGBT IRG4PC40UDPBF V _{ce} =600, I _c =20A
D1-D5	RHEP30120 V _{rr} =1200V, I=30A
L1	2.2mH
Lf	3mH
C1-C2	220μF V _{dc} =500V
Switching Frequency	20KHz
Sampling Frequency	78KHz

V. SIMULATION RESULTS

In order to verify that the proposed inverter can be practically implemented in a PV system, simulations were performed by using MATLAB SIMULINK. It also helps to confirm the PWM switching strategy which then can be implemented. It consists of two reference signals and a triangular carrier signal. Both the reference signals are compared with the triangular carrier signal to produce PWM switching signals for switches S1–S5. Note that one leg of the inverter is operating at a high switching rate equivalent to the frequency of the carrier signal, whereas the other leg is operating at the rate of fundamental frequency (i.e., 50 Hz). The switch at the auxiliary circuit S1 also operates at the rate of the carrier signal. As mentioned earlier, the modulation index M will determine the shape of the inverter output voltage V_{inv} and the grid current I_g . Fig.7 shows V_{inv} and I_g for different values of M .

The dc-bus voltage is set at 400 V ($> \sqrt{2}V_g$; in this case, V_g is 240 V) in order to inject current into the grid. Fig. 7(a) shows that V_{inv} is less than $\sqrt{2}V_g$ due to M being less than 0.5. The inverter should not operate at this condition because the current will be injected from the grid into the inverter, rather than the PV system injecting the current into the grid, as shown in Fig. 7(b). Over modulation condition, which happens when $M > 1.0$, is shown in Fig. 7(c). It has a flat top at the peak of the positive and negative cycles because both the reference signals exceed the maximum amplitude of the carrier signal. This will cause I_g to have a flat portion at the peak of the sine waveform, as shown in Fig. 7(d). To optimize the power transferred from PV arrays to the grid, it is recommended to operate at $0.5 \leq M \leq 1.0$. V_{inv} and I_g for optimal operating condition are shown in Fig. 7(e) and (f), respectively. As I_g is almost a pure sine wave, the THD can be reduced compared with that under other values of M .

VI. CONCLUSION

This paper presented a single-phase multilevel inverter for PV application. It utilizes two reference signals and a carrier signal to generate PWM switching signals. The circuit topology, modulation law, and operational principle of the proposed inverter were analyzed in detail. A FUZZY control is implemented to optimize the performance of the inverter. MATLAB/SIMULINK results indicate that the THD of the Fuzzy Controller Circuit is much lesser than that of the THD of the PI Controller Circuit. Furthermore, both the grid voltage and the grid current are in phase at near-unity power factor.

REFERENCES

- [1]. Xiong Liu, Peng Wang and Poh Chiang Loh "A Hybrid AC/DC Microgrid and Its Coordination Control," *IEEE Trans.*
- [2]. M. E. Ropp and S. Gonzalez, "Development of a MATLAB/simulink model of a single-phase grid-connected photovoltaic system," *IEEE Trans. Energy Conv.*, vol. 24, no. 1, pp. 195–202, Mar. 2009.
- [3]. K. H. Chao, C. J. Li, and S. H. Ho, "Modeling and fault simulation of photovoltaic generation systems using circuit-based model," in *Proc. IEEE Int. Conf. Sustainable Energy Technol.*, Nov. 2008, pp. 290–294.
- [4]. J. M. Carrasco, L. G. Franquelo, J. T. Bialasiewicz, E. Galvan, R. C. PortilloGuisado, M. A. M. Prats, J. I. Leon, and N. Moreno-Alfonso, "Power-electronic systems for the grid integration of renewable energy sources: A survey," *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1002–1016, Aug. 2006.
- [5]. V. G. Agelidis, D. M. Baker, W. B. Lawrance, and C. V. Nayar, "A multilevel PWM inverter topology for photovoltaic applications," in *Proc. IEEE ISIE*, Guimarães, Portugal, 1997, pp. 589–594.
- [6]. S. Kouro, J. Rebolledo, and J. Rodriguez, "Reduced switching-frequency modulation algorithm for high-power multilevel inverters," *IEEE Trans. Ind. Electron.*, vol. 54, no. 5, pp. 2894–2901, Oct. 2007.
- [7]. S. J. Park, F. S. Kang, M. H. Lee, and C. U. Kim, "A new single-phase five level PWM inverter employing a deadbeat control scheme," *IEEE Trans. Power Electron.*, vol. 18, no. 18, pp. 831–843, May 2003.
- [8]. L. M. Tolbert and T. G. Habetler, "Novel multilevel inverter carrier-based PWM method," *IEEE Trans. Ind. Appl.*, vol. 35, no. 5, pp. 1098–1107, Sep./Oct. 1999.
- [9]. CHUEN CHIEN LEE "Fuzzy Logic in Control System: Fuzzy Logic Controller-Part 1," *IEEE Trans.*
- [10]. A. Nabae and H. Akagi, "A new neutral-point clamped PWM inverter," *IEEE Trans. Ind. Appl.*, vol. IA-17, no. 5, pp. 518–523, Sep./Oct. 1981.
- [11]. J. Pou, R. Pindado, and D. Boroyevich, "Voltage-balance limits in four level diode-clamped converters with passive front ends," *IEEE Trans. Ind. Electron.*, vol. 52, no. 1, pp. 190–196, Feb. 2005.
- [12]. S. Alepuz, S. Busquets-Monge, J. Bordonau, J. Gago, D. Gonzalez, and J. Balcells, "Interfacing renewable energy sources to the utility grid using a three-level inverter," *IEEE Trans. Ind. Electron.*, vol. 53, no. 5, pp. 1504–1511, Oct. 2006.
- [13]. T. Meynard and H. Foch, "Multi-level choppers for high voltage applications," *Eur. Power Electron. J.*, vol. 2, no. 1, pp. 45–50, Mar. 1992.

Application of Taguchi Methodology in Selection of Process Parameters For Induction Hardening Of EN8 D Steel

P.G.Kochure¹, K.N.Nandurkar²

Department of Production Engineering, K.K.Wagh Institute of Engineering Edu. And Research, Nashik (India)

ABSTRACT: In this paper, Taguchi method is applied for selection of best working parameters for induction hardening of EN8 D steel. A orthogonal array L9, Signal to Noise ratio, analysis of variance (ANOVA) are applied to study, performance characteristics of induction hardening process. Hardness, case depth, has been considered as performance characteristics. An analysis of variance (ANOVA) of response variables shows a significant influence of process variable power, heating time. The experimental results shows that the predicted regression models suggested could describe the performance characteristics within the limits of the factors being investigated. The results have been verified by confirmation tests.

Keywords: Anova, Induction Hardening, Orthogonal array, Regression models, Signal to noise ratio, Taguchi method.

I. INTRODUCTION

Economic globalization and rapid and continuous appearing of new technologies, mobilized organizations to obtain the maximum degrees of competitiveness, high quality products in short time in order to ensure their survival and growth in the market. In this scenario, organizations started to look for quality not only in their products but also in its production processes. Many surface hardening (strengthening) treatments are required corresponding variations in physical, chemical and mechanical properties, and re-arrangements of atoms in metals and their alloys. An induction hardening is one of the important treatment among them. Induction hardening process have found ever-increasing applications to improve the performance and life of the parts used in automobile engineering. Thin surface layers i.e 0.25 to 2.3 mm of the work piece made of steel can be hardened by this process. Y.Totic, R.Sadeler, H.Altum investigated the effects of heating time, feed rate and temperature on wear characteristics of AISI 4140 steel in induction hardening process.[13] Julie.K, Timothy James.[12] studied the effect of feed rate and gap between coil and work piece, quench distance and part temperature by using design of experiment neural network optimization technique on induction hardening process and reported a significant improvement in the process. R.Kolleck, R.Veit[7] focused on reduction of processes cycle time, rising energy, costs eco friendly process and need of new heating technologies in hardening process and proposes the inductive heating alternative methodology for boron alloyed steel. Robert Cryderman, Nima Sham Saei, Al Fatemi [9] In this paper study investigates the influences of induction hardened parts produced from steel bars. Resit Unal, Edwin B.Dean[8] In this paper authors were presented the overview of the

Taguchi method its steps involved and state that, it is a systematic and efficient approach for determining the optimum experimental configuration of design parameters for performance, quality and cost. Principal benefits include considerable resource savings determination of important factors affecting operation, performance and cost, and quantitative recommendations for design parameters which achieve lowest cost, high quality solutions. In this paper, best working parameters are selected for generation of desired hardness values and pattern in EN8 D steel, by using Taguchi method. Because Taguchi method showed to be a very useful in process improvement provide confident information about influence of factors on a response variable and less number of experiments than traditional method to improve the process while not compromising the desired goals.[3] Taguchi method based design of experiments has been used to study the effects of two process parameters i.e. power and traverse speed.(scanning speed) on important output parameters. Taguchi approach provides a new experimental strategy in which a modified and standardized form of design of experiment (DOE) is used. This technique helps to study effect of many factors (variables) on the desired quality characteristic most economically. By studying the effect of individual factors on the results, the best factor combination can be determined [6]. Taguchi designs experiments using specially constructed tables known as "orthogonal array" (OA). The use of these tables makes the design of experiments very easy and consistent and it requires relatively lesser number of experimental trials to study the entire parameter space. As a result, time, cost, and labour saving can be achieved. The experimental results are then transformed into a signal-to-noise (S/N) ratio. Taguchi recommends the use of the S/N ratio to measure the quality characteristics deviating from the desired values. Usually, there are three categories of quality characteristic in the analysis of the S/N ratio, i.e. The-lower-the-better, the-higher-the-better, and the nominal-the-better. The S/N ratio for each level of process parameters is computed based on the S/N analysis. Regardless of the category of the quality characteristic, a greater S/N ratio corresponds to better quality characteristics. Therefore, the optimal level of the process parameters is the level with the greatest S/N ratio. Furthermore, a statistical analysis of variance (ANOVA) is performed to see which process parameters are statistically significant. With the S/N and ANOVA analyses, the optimal combination of the process parameters can be predicted. Finally, a confirmation experiment is conducted to verify the optimal process parameters obtained from the parameter design. [10]

II. EXPERIMENTAL PROCEDURE.

2.1 . Experimental apparatus

High frequency induction hardening machine (40 kw, 200khz) make sanket power systems has been used for the performing the experiments . A source of high frequency of electricity is used to drive a large alternating current through a copper coil. The passage of a current through this coil generates a intense and rapidly changing magnetic field in the space within the work coil. The work piece to be heated is placed within this magnetic field where eddy current is generated within the work piece and resistance leads to heating of work piece. The core of the work piece remains unaffected by this process . Induction hardening temperature was above 850 °C .

2.2 Work piece material



Fig. 1a : case study shaft

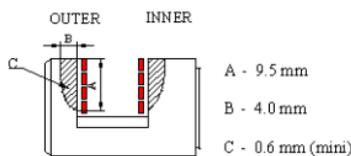


Fig. 1 b :detailed view of slot

The automobile shaft i.e. selector and shifting shaft was selected for study. The work piece material used for present study was EN8 D steel. Its composition was 0.40% Carbon, 0.22% Silicon, 0.71% Manganese , 0.02% Sulphur , and 0.024% Phosphorus . Material conforms to BS 970(1955) for EN8 D steel. This material is suitable for variety of automotive components such as axle, crank shaft, spline shafts and gears etc. The length of work piece is 177 and diameter 19mm respectively. The static induction hardening has been performed on the slot portion of work piece.

2.3 Experimental plan.

In this investigation two factors at three levels have been studied. Factors and their levels are shown in table 1 . The factors and levels are selected on the basis of literature review and their range were finalized after pilot runs.[11]

Table 1 : process parameters and their levels

Symbol	Process parameters	Level I	Level II	Level III
P	Power (kw)	10	12	14
T	Heating time (sec)	02	03	04

The values of factors of induction hardening process were utilized for conducting design of experiments in induction hardening machine for EN8 D steel. The

response variables to be investigated were hardness HRC, and its developed pattern i.e. case depth , Based on available in put parameters and their levels an orthogonal array lay out is designed for Taguchi method by using soft ware MINITAB .

Table-2 – Process Parameters and their levels

Exp. Runs	Power	Heating time	Hardness	Case depth(mm)					
				Outer			Inner		
No	kw	sec	HRC	V	T	C	V	T	C
1	10	02	55	8.70	2.0	0.4	8.75	2.2	0.5
2	10	03	56	8.75	2.3	0.5	8.85	2.4	0.6
3	10	04	57	9.20	2.5	0.8	9.25	2.7	0.9
4	12	02	57	9.00	2.2	0.5	9.05	2.3	0.6
5	12	03	57	9.20	2.4	0.9	9.25	2.8	0.7
6	12	04	58	9.45	3.3	1.5	9.50	3.6	1.6
7	14	02	58	9.00	3.0	1.0	9.10	3.2	1.1
8	14	03	58	9.35	3.2	1.3	9.30	3.5	1.4
9	14	04	60	9.60	4.0	1.8	9.50	3.9	1.8

V- Vertical , T - Top, C – Centre

2.4 Experimental technique

As shown in table 2 nine experiments are conducted. Each experiment is conducted for three times to reduces the errors. Hardness was measured by Rockwell hardness testing machine for C scale at 150 kg load, having diamond indenter at 120 degrees . While automatic Vickers’s hardness testing machine is used to measure case depth achieved at various locations.

III. RESULTS AND DISCUSSIONS

The mean of three experimental values of each run is mentioned in the table 2 further analysis of results were carried by using software MINITAB.

3.1 Analysis of variance (ANOVA)

The ANOVA tables has been used to investigate and test for significance of design parameters. It indicates which parameters are significantly affecting the output parameters. In the analysis the sum of squares and variance are calculated . F-test values at 95% confidence level is used to decide the significant factors affecting the process and percentage contribution. Since the p values are less than 0.05 and larger F values indicates that these factors have statically significant effects on the performance . The ANOVA analysis for response are shown in table 3.

Table 3: ANOVA results and percentage contribution for responses.[2][5]

Source	DF	SS	MS	F	% contribution
Heating time	2	0.80222	0.401111	11.65	43.704
Error	4	0.13778	0.034444		7.507
Total	8	1.83556			100
S = 0.1856 R-Sq = 92.49% R-Sq(adj) = 84.99%					
All F-ratios are based on the residual mean square error.					
For Hardness					
Power	2	10.6667	5.33333	32.00	66.667
Heating time	2	4.6667	2.33333	14.00	29.167
Error	4	0.6667	0.16667		4.166
Total	8	16.0000			100
S = 0.4082 R-Sq = 95.83% R-Sq(adj) = 91.67%					
For Case depth					
Outer vertical					
Power	2	0.317222	0.158611	20.39	42.171
Heating time	2	0.403889	0.201944	25.96	53.693
Error	4	0.031111	0.007777		4.136
Total	8	0.752222			100
S = 0.08819 R-Sq = 95.86% R-Sq(adj) = 91.73%					

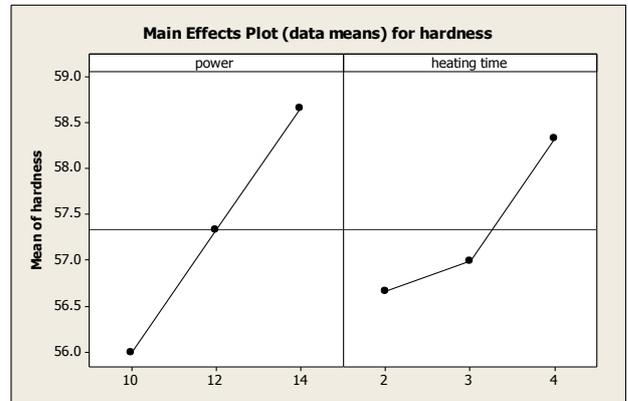


Fig. 2 a : plot for hardness

Outer top					
Power	2	2.00667	1.00333	24.08	59.369
Heating time	2	1.20667	0.60333	14.48	35.700
Error	4	0.16667	0.04167		4.931
Total	8	3.38000			100
S = 0.2041 R-Sq = 95.07% R-Sq(adj) = 90.14%					
Outer centre					
Power	2	0.96000	0.48000	20.57	51.064
Heating time	2	0.82667	0.41333	17.71	43.972
Error	4	0.09333	0.02333		4.964
Total	8	1.88000			100
S = 0.1528 R-Sq = 95.04% R-Sq(adj) = 90.07%					
Power	2	0.215000	0.107500	32.25	41.747
Heating time	2	0.286667	0.143333	43.00	55.663
Error	4	0.013333	0.003333		2.590
Total	8	0.515000			100
S = 0.05774 R-Sq = 97.41% R-Sq(adj) = 94.82%					
Inner top					
Power	2	1.82889	0.914444	20.57	59.724
Heating time	2	1.05556	0.527777	11.87	34.470
Error	4	0.17778	0.044444		5.806
Total	8	3.06222			100
S = 0.2108 R-Sq = 94.19% R-Sq(adj) = 88.39%					
Inner centre					
Power	2	0.89556	0.447778	13.0	48.78

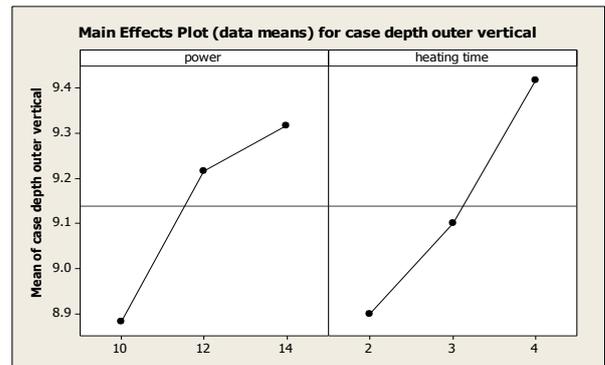


Fig. 2 b: plot for case depth outer vertical

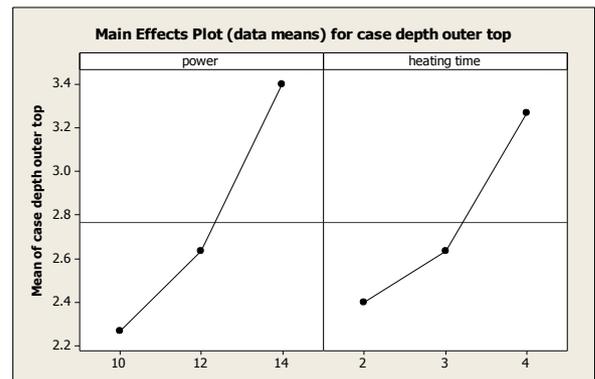


Fig. 2 c : plot for outer top

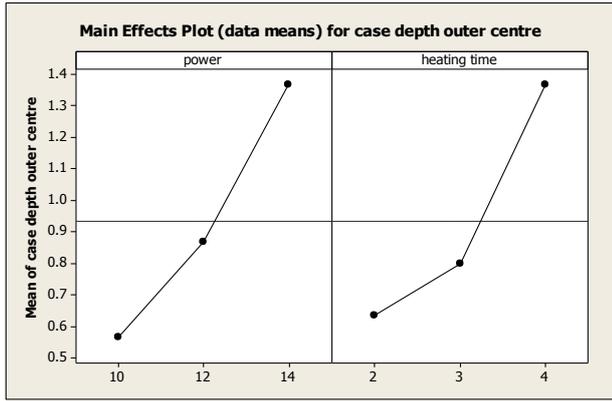


Fig. 2 d : plot for outer centre

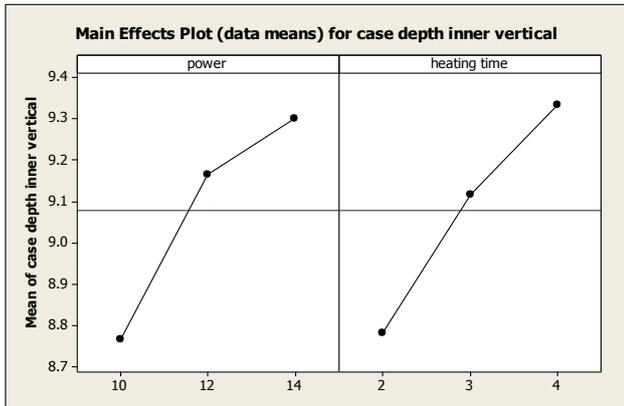


Fig. 2 e : plot for case depth inner vertical

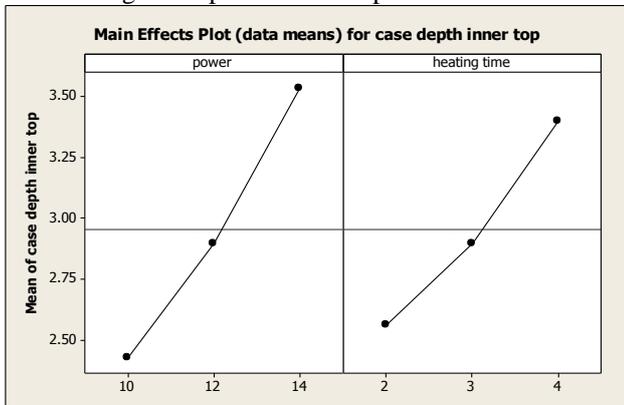


Fig. 2 f: plot for case depth inner top

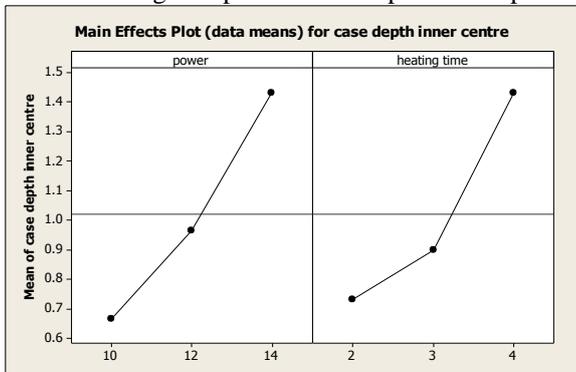


Fig. 2 g: plot for case depth inner centre

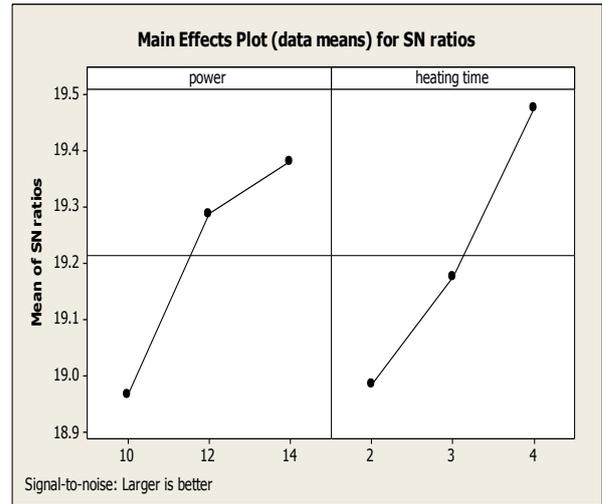


Fig. 3 a : plot for hardness

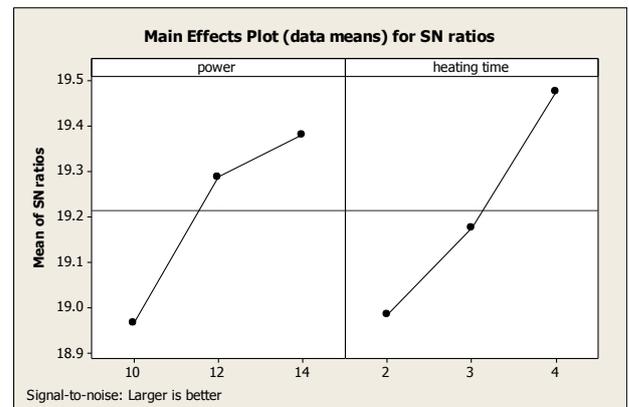


Fig. 3 b: plot for case depth outer vertical

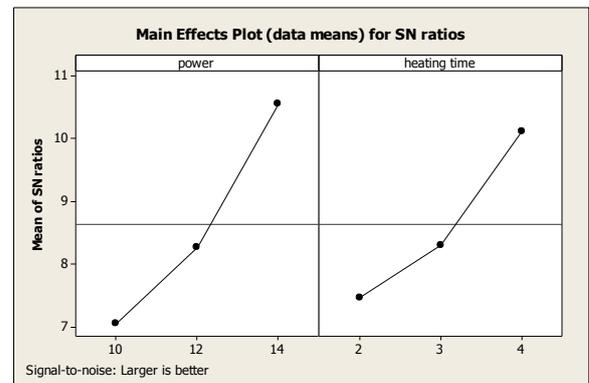


Fig. 3 c : plot for outer top

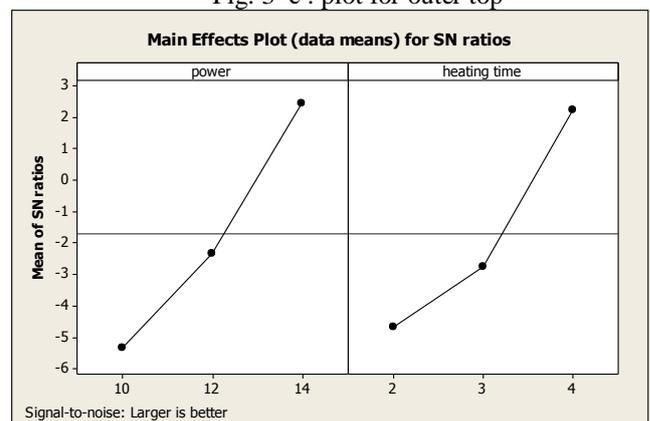


Fig. 3 d : plot for outer centre

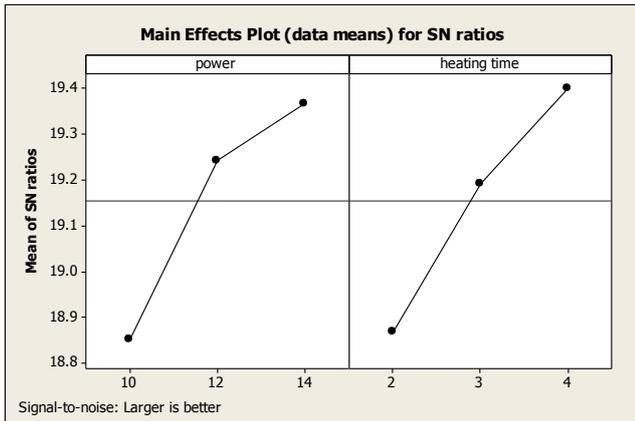


Fig. 3 e : plot for case depth inner vertical

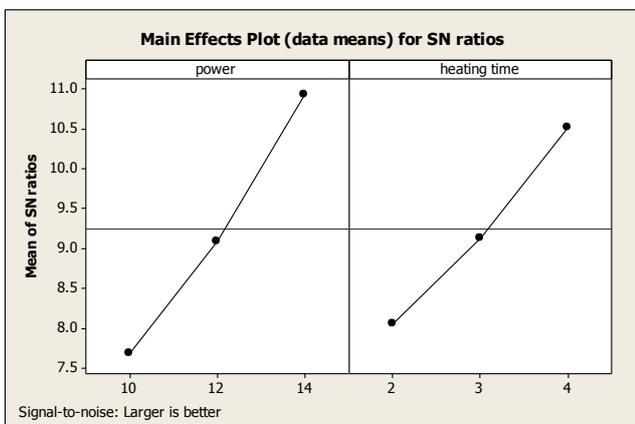


Fig. 3 f: plot for case depth inner top

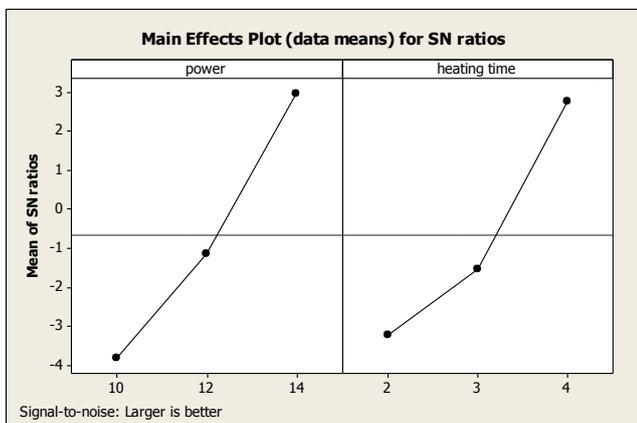


Fig. 3 g: plot for case depth inner centre

Figure 2a and 2b-2g shows main effects plot (data means) of mean for hardness and case depth at various locations i.e. outer vertical, top, centre and inner vertical, top, centre respectively. Figure 3 a – 3g shows S/N ratio graphs for hardness and corresponding case depth. Where the horizontal line is the value of total mean of S/N ratio. Basically larger the S/N ratio better is the quality characteristics for process. As per the S/N ratio analysis from the graphs the levels of parameter to be set getting optimum values of hardness and case depth are P3-T3 i.e. 14kw-4sec According to ANOVA analysis as shown in the table 3 most effective parameter with respect to hardness

and case depth is power. It is also observed that in outer vertical and inner vertical case depth pattern achievement heating time plays a important role. The analysis of graphs and ANOVA results clearly indicates that increase in power and heating time increases hardness and corresponding case depth at various locations of work piece. Table 4: Significance of induction hardening parameters for hardness and case depths.

Process parameter	Mean S/N ratio			Significance of induction hardening parameter Max-Min
	Level I	Level II	Level III	
Hardness				
P	34.96	35.17	35.37	0.41
T	35.06	35.12	35.32	0.25
Case depth				
Outer top				
P	7.071	8.274	10.262	3.491
T	7.470	8.310	10.123	2.653
Outer centre				
P	-5.306	-2.312	2.461	7.767
T	-4.660	-2.726	2.230	6.889
Inner centre				
P	-3.791	-1.151	2.952	6.743
T	-3.210	-1.537	2.758	5.967
Outer vertical				
P	18.97	19.31	19.38	0.41
T	18.99	19.19	19.48	0.49
Inner vertical				
P	19.03	19.34	19.37	0.33
T	19.05	19.21	19.48	0.41

Significance of induction hardening parameters (difference between max and min values) indicates that power is significantly contributing towards the induction hardening performance as difference gives higher values only in case depth outer vertical and case depth inner vertical time was more influential from above study finds that optimized levels of parameters are P3-T3 i.e. 14 kw-4 sec.

Confirmation test-

After identifying the best levels of process parameters, a new experiment was designed and conducted with predicted levels of process parameters optimum parameters found are P-14kw, T-4 sec, comparison between estimated values and confirmation test value shows a small difference between these values. These values correlate each other. This indicates that ANOVA results are closely match with the Taguchi results.[1]

IV. REGRESSION MODEL.

Multiple linear regression equations were modelled for relationship between process parameters i.e. power and heating time to evaluate hardness and case depths for any combination of factor levels in a range specified model for multiple regression equation is

$$y = \beta_0 + \beta_1 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_p x_p + \varepsilon \quad (1)$$

Where y is dependent parameter, $x_1, x_2, x_3, \dots, x_p$ are independent parameters, β are regression parameters and ε is residue.[3]

The regression coefficients have been obtained by using experimental data. The regression equations for response characteristics as a function of two parameters of the material EN8 D steel considered in this experiment are given below. The insignificant coefficients identified from ANOVA have been omitted from equations for various responses. In this case, regression equations are formulated in terms of parameters i.e. power and heating time by using MINITAB software.

The regression model equations for response in terms of power and heating time are given as below

- Hardness = 46.8 + 0.667 Power + 0.833 Heating time (2)

- Case depth outer vertical = 7.07 + 0.108 Power + 0.258 Heating time (3.1)

- Case depth outer top = -1.93 + 0.283 Power + 0.433 Heating time (3.2)

- Case depth outer centre = -2.53 + 0.200 Power + 0.367 Heating time (3.3)

- Case depth inner vertical = 7.45 + 0.0875 Power + 0.225 Heating time (3.4)

- Case depth inner top = -1.59 + 0.257 Power + 0.417 Heating time (3.5)

- Case depth inner centre = -2.33 + 0.192 Power + 0.350 Heating time (3.6)

•
Confirmation for regression equations

Expected values given by models of hardness and case depth at various locations are obtained in any combination of factor levels as specified in table 2. Percentage variation estimated for this equation 2 for hardness in range of 0.055% to 3.161%. Where as for the regression equations 3.1 to 3.6 case depth at various locations i.e. outer vertical, top, centre as well as inner vertical, top, centre are calculated as 0.001% to 14.1%. These results obtained by regression equations closely co relate with each other which validated the regression equations developed. Thus the developed equations can be used to predict hardness and case depth values at different locations such as outer vertical, top, centre and inner vertical top and centre respectively for any combination of factor levels in the specified range[2].

V. CONCLUSION.

Taguchi method of experimental design with L9 orthogonal array has been applied for selection of optimum process parameters of induction hardening of EN8 D steel. The experimental investigation shows the effects of process parameters such as power, heating time on hardness and

case depth pattern achieved on work piece. The optimum parameters found are 14 kw power and heating time 4 sec. power is the most influential parameter. Further multiple regression equations are formulated for estimating predicted values of hardness and case depths at various locations such as case depths at outer & inner vertical, top and centre portion of slots for a specified range. The results obtained by regression equations closely co relate each other which validates the regression equation developed.

VI. ACKNOWLEDGMENT.

The authors wish to express their deep sense of gratitude to M/S Industrial Heat Treater, Nashik for providing induction hardening and metallurgical laboratory facilities. We also wish to acknowledge the technical support and encouragement provided by Mr. Vinit Pol CEO and their technical staff as well as faculty of production engineering department of KKWIEE&R Nashik and Mr. Rajesh Deshpande, Mr. Yogesh Joshi, Directors of A. R. Technology Satpur Nashik for their enthusiasms to the valuable, unprofitable contribution on fulfilling this paper.

REFERENCES .

- [1] Amit kohli and Hari Singh, "Optimization of process parameters in induction hardening using surface response method" Indian Academy sciences ,sadhana Vol 36 –Part 2, 2010 P 87-91
- [2] Bala Murugan Gopalsamy, Biswanath Mondal and sukamal ghosh, "Taguchi method and ANOVA : An approach for process parameters optimization of hard machining while machining hardened steel" Journal of Scientific & Industrial Research Vol 68, August 2009, p 686-695.
- [3] Douglas C. Montgomery, Wiley India, Pvt. Ltd. New Delhi "Design and analysis of experiments" 2009, P3-13,19,392-400
- [4] Keith M. Bower M.S. MINITAB Inc. "ANOVA by MINITAB" Scientific Computing and Instrumentation, Feb.2000
- [5] M .N .Das, N. C. Giri, New Age international (p) Ltd. Publisher, New Delhi "Design and Analysis of experiments" 1999,P 13-15
- [6] M. Kaladhar, K.V.Subbaiah, Ch.Srinivasa Rao, And k.narayana Rao, "Application of Taguchi approach and Utility Concept in solving the Multi-objective Problem when turning AISI 202 Austenitic Stainless Steel" Journal of Engg. Science and Technology, 2011, p 55-61
- [7] R. Kolleck (3)a,*, R. Veit a, M. Merklein b, J. Lechler b, M. Geiger (1)b "Investigation on induction heating for hot stamping of boron alloyed steels" 2009
- [8] Resit Unal, Edwin B.Dean, "Taguchi approach to design optimization for quality and cost : An overview" Annual Conference of the international Society of parametric Analysts, 1991 p 1-9
- [9] Robert Crydermana, Nima Shamsaeib, Ali Fatemib, "Effects of continuous cast section size on torsion cook deformation and fatigue of induction hardened 1050 steel shafts" 2010
- [10] S. Kamaruddin, Zahid A. Khan and S. H. Foong,

“ Application of Taguchi method in the optimization of injection moulding parameters for manufacturing products from plastic blend”. IACSIT International Journal of Engineering and Technology, Vol.2, No.6, December 2010 , p 574-580.

- [11] Software MINITAB release 14
- [12] Timothy James Stich, Julie K., “ The Application of Artificial Neural Networks to Monitoring and Control of an Induction Hardening Process. ” Journal of Industrial Technology, **16**(1):1-11(2000).
- [13] Y. Totik, R. Sadeler, H. Altun and M. Gavali. “ The effects of Induction Hardening on Wear Properties of AISI 4140 Steel in dry sliding conditions ”, Materials & Design **24**(1):25-30(2003).

Reducing Real Time Service Delays Using Map reduce Frame Work

K.M.DhanaSailaja¹, E.V.Prasad²

(M.Tech, University College of Engineering JNTUK, Andhra Pradesh, India)
(Professor & Registrar, University College of Engineering JNTUK, Andhra Pradesh, India)

ABSTRACT: Recently the demand of real-time data services has increased in various applications such as manufacturing, Web-servers, and e-commerce and they are becoming sophisticated in their real-time data needs. Real time data services has to provide better QoS parameters because end users may ignore the service when the database service delay is high and temporally inconsistent data is available. Due to dynamic workloads providing better QoS in data services is a challenging issue. In this paper to overcome this problem we are using map-reduce framework to estimate the real time service delays. MapReduce framework process large amount of data in a parallel way. It is gaining lot of interest in data mining, because the programmer is abstracted to the data storage, distribution, replication, load balancing and it uses functional programming. It has two functions map and reduce. Several experiments have been conducted on various data sets to calculate the performance of the proposed technique.

Keywords : Web-servers, Map Reduce, e-commerce, Data mining, Quality of service

I. INTRODUCTION:

Recently the demand of real-time data services has increased in various applications such as manufacturing, Web-servers, and e-commerce and they are becoming sophisticated in their real-time data needs [1], [2]. In the past decade demand is increasing provisioning for quality of service (QoS) guarantees to various network applications and clients. The data normally span from low-level control data, typically acquired from sensors, to high-level management and business data. In these applications, it is desirable to process user requests within the timeline using fresh data. In dynamic systems, such as Web servers and sensor networks with non-uniform access patterns, the workload of real-time databases (RTDB) cannot be predicted accurately. Hence, the RTDBs can be overloaded, uncontrolled deadline misses and data freshness may not be possible during the transient overloads. To provide service quality we propose a quality of service (QoS) sensitive method that provides a set of requirements to improve the performance of the database even in the unpredictable workloads conditions. In some applications like Web service it is desirable that the QoS does not vary significantly from one transaction to another. It is emphasized that individual QoS needs requested by transactions are enforced and any deviations from the QoS needs should be uniformly distributed among the clients to ensure QoS fairness. Imprecise computation techniques [3] have been proposed by various authors to allow flexibility and for achieving graceful degradation during transient

overloads. These techniques make it possible to trade off resource needs for the quality of a requested service and they have been successfully applied in timeliness is emphasized applications where a certain degree of imprecision can be tolerated [4], [5], [6]. Real time data services has to provide better QoS parameters because end users may ignore the service when the database service delay is high and temporally inconsistent data is available. Due to dynamic workloads providing better QoS in data services is a challenging issue. In this paper to overcome this problem we are using map-reduce framework to estimate the real time service delays. Our MapReduce framework process large amount of data in a parallel way and it is gaining lot of interest in data mining, because the programmer is abstracted to the data storage, distribution, replication, load balancing and it uses functional programming. It has two functions map and reduce. In the map stage passes the data over the input file and outputs (key, value) pairs, the shuffling stage transfers the mappers output to the reducers based on the key and finally the reducer processes the received pairs and outputs the final result. Map Reduce is a useful tool for large amount of data analysis because of its scalability, simplicity and low cost to build large clouds of computers.

The rest of the paper is organized as section 2: discuss about the related work, section 3: presents the Proposed Solution, section 4: discuss about Map Reduce, section 5: discuss about Experimental setup, section 6: concludes the paper.

II. RELATED WORK:

In e-commerce, the real time data service quality is mainly determined by both networks and data transfer rate between disks. The quality of service is mainly depending on the whole data that contain embedded objects. All the existing techniques, measure service quality with respect to a single packet in networks [11], [13], [15] or an individual request [7], [9], [12], [17], [21] or connection [20], [22], [24] in Web servers. C. Dovrolis [15], proposed a technique to provide different QoS levels between multiple aggregated traffic classes within a network. T.F.

Abdelzaher [7], bound the server-side delay of individual client request. This algorithm mainly focuses on providing differentiated services to different client classes using priority-based scheduling [14], [16]. The aim of the algorithm is to provide better services to the premium class than to the basic class by adjusting the priority of the allocated processes between the classes on either user level or kernel level. Existing techniques are not providing any guarantee to QoS. To overcome this problem authors proposed queuing-theoretic approaches. It is well known that the delay upper bound in a G/G/1 is determined by the

system load and the variance of user's requests inter arrival and service time distributions. B. Urgaonkar [24] proposed an approach in which the resource allocation is controlled to adjust the load of a class so that the delay equals the upper bound and the performance of the approach is highly depend on the variance estimation, which is very difficult to measure accurately in the dynamic workload cases. Traditional linear feedback control is used to control the resource allocation in Web servers [7], [21]. Because the behavior of a Web server changes continuously, the performance of the linear feedback control is limited.

To overcome the problem of lack of accurate server model A. Kamra and M. Karlsson [17, 18] the parameters of the controllers were adjusted based on the workload. V. Sundaram [23] presented an approach where the resource allocation was controlled based on the past allocations of delivered service quality. All these approaches provide better performance than nonadaptive linear feedback control approaches under dynamic workload situations. Later B. Li [19] presented a fuzzy control model to address the nonlinear QoS requirements of different multimedia applications under different resource constraints. In [11], author proposed a set of rules to dynamically adjust the target delay ratios between various traffic flows to reduce the effect of bursty traffic in the basic class on the delay of the premium class.

III. PROPOSED SOLUTION:

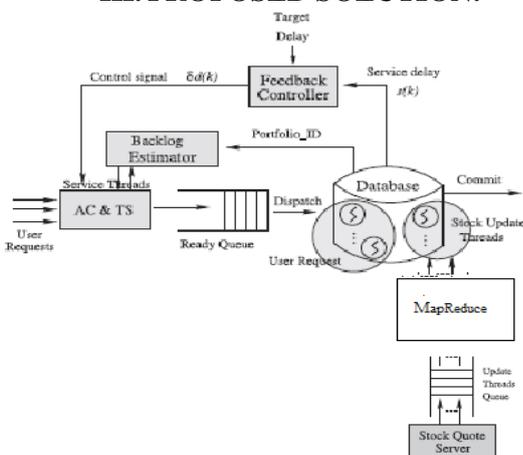


Fig 1: Proposed architecture

Figure 1 shows the proposed architecture of the database system. In this architecture we added the Map Reduce to the Chronos architecture [29] to provide better QoS in real time data services. It has database backlog estimator, admission controller, traffic smoother, feedback controller, MapReduce and a database server block. The feedback controller is designed based on either linear control theory or fuzzy logic control. The database server processes service requests and updates stock prices periodically received from the stock quote server to provide the updated stock prices. We consider periodic temporal updates that are commonly used in RTDBs for data temporal consistency [30] and a fixed time interval is selected for updating each stock price in a range [0.2 s, 5 s].

A periodic temporal updates are rarely considered in RTDBs due to the difficulties for defining and

maintaining the notion of temporal consistency, backlog estimation for periodic updates and estimating the database backlog to service user data service requests. The dedicated update threads are scheduled in a separate queue ahead of user requests and they have to wait for next available threads. High priority is assigned to temporal data updates to provide data freshness. Map Reduce process large amount of data in a parallel way and it has two functions map and reduce. In the map stage passes the data over the input file and outputs (key, value) pairs, the shuffling stage transfers the mappers output to the reducers based on the key and finally the reducer processes the received pairs and outputs the final result.

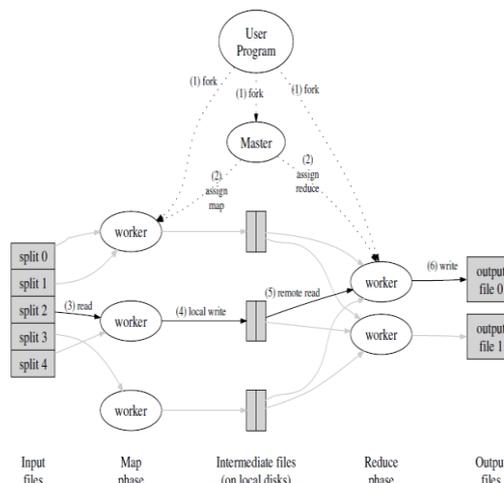


Fig 2: Block diagram of MapReduce

IV. MAPREDUCE:

Map Reduce framework process large amount of data in a parallel way and it is gaining lot of interest in data mining, because the programmer is abstracted to the data storage, distribution, replication, load balancing and it uses functional programming. It has two functions map and reduce. In the map stage passes the data over the input file and outputs (key, value) pairs, the shuffling stage transfers the mappers output to the reducers based on the key and finally the reducer processes the received pairs and outputs the final result. Map Reduce is a useful tool for large amount of data analysis because of its scalability, simplicity and low cost to build large clouds of computers.

1.1. Programming Model

The Map Reduce takes a set of input key/value pairs and produces a set of output key/value pairs. It has two functions Map and Reduce. Map function written by the user takes an input pair and produces a set of intermediate key/value pairs. The MapReduce library groups all intermediate values with same intermediate key and passes them to the Reduce function. Further Reduce function written by the user accepts an intermediate key and a set of values for that key. It merges together these values to form a possibly smaller set of values and zero or one output value is produced per Reduce invocation. Using an iterator the intermediate values are supplied to the reduce

function which allows to handle large values to fit into the memory.

1.2. Example

Consider the problem of counting the number of occurrences of each word in a large collection of documents. The user would write code similar to the following pseudo-code:

```
map(String key, String value):
    // key: document name
    // value: document contents
    for each word k in value:
        EmitIntermediate(k, "1");

reduce(String key, Iterator values):
    // key: a word
    // values: a list of counts
    int res = 0;
    for each j in values:
        res += ParseInt(v);
    Emit(AsString(res));
```

The map function ejects each word plus an associated count of occurrences. The reduce function adds all counts ejected for a particular word.

1.3. Types

The map and reduce functions called by the user have associated types:

```
map (k1,v1) = list(k2,v2)
reduce (k2,list(v2)) = list(v2)
```

The input keys and values are drawn from a different domain than the output keys and value and the intermediate keys and values are from the same domain as the output keys and values.

The main problem in MapReduce for clustering very large datasets is how to minimize the I/O cost and how to minimize the network cost among processing nodes. To overcome these problems we used the ParC (Parallel Clustering) and SnI (Sample-and-Ignore) method. The ParC [27] algorithm partition the input data and assign each partition to an individual system then each system group the data partition in to a cluster, named as β - clusters, and finally merge the β -clusters to get the final clusters. This algorithm reads the dataset once to minimize disk access which is the common way used by serial algorithms to decrease the computational costs. But it does not discuss how to minimize the network traffic [28]. To address this problem we used SnI method, it samples the input data and creates an initial set of clusters and later the input data is filtered to include unclassified elements. Finally the clusters found by the reducers are merged with the clusters from the sampling phase using the same merging strategies used in ParC.

V. EXPERIMENTAL SETUP:

To evaluate the performance of the proposed algorithm, we use three machines each of them is same configuration such as the dual core 1.6GHz CPU and 1GB memory with Linux operating system. A Chronos server,

clients and a stock quote server run on each of them, respectively.

For 80% of time, a client issues a query about stock prices. In the remaining time client requests a portfolio browsing, purchase or sale transaction at a time. Most data service requests in e-commerce are queries. At the beginning of the experiment, the inter-request time (IRT) is randomly distributed in [3s, 4s]. At 200s, the range of the IRT is suddenly reduced to model heavy workload changes, and stays in the new range until the end of the experiment at 500s.

For performance comparisons, we consider three approaches shown in Table 1. Open is the basic Berkeley DB [26] without any control facility. AC model control the incoming transactions in proportion to the service delay error under overload. FC-Q finds the relation between the queue size and response time, similar to [25].

Table 1: Tested Approaches

OPEN	Pure Berkeley
AC	Ad-hoc admission control
FC-Q	Feedback control (FC)
MapReduce	MapReduce control

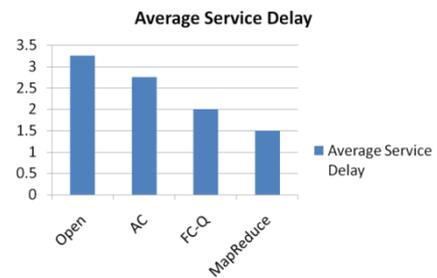


Fig 3: Average Service Delay

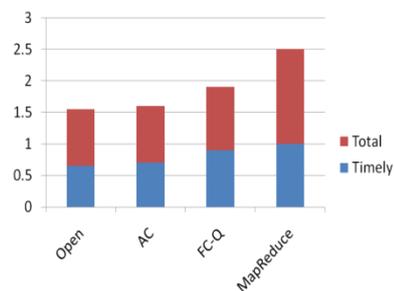


Fig 4: Number of Data processed

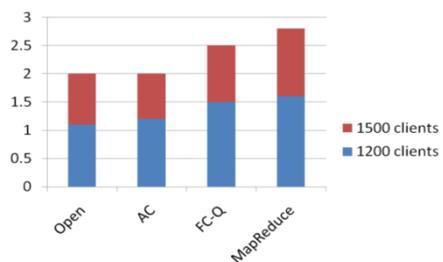


Fig 5: Number of Data processed

VI. CONCLUSION:

Real time data services has to provide better QoS parameters because end users may ignore the service when the database service delay is high and temporally inconsistent data is available. Due to dynamic workloads providing better QoS in data services is a challenging issue. In this paper to overcome this problem we are using map-reduce framework to estimate the real time service delays. MapReduce framework process large amount of data in a parallel way. It is gaining lot of interest in data mining, because the programmer is abstracted to the data storage, distribution, replication, load balancing and it uses functional programming. It has two functions map and reduce. Experimental results relived that our proposed approach support the desired average/transient data service delay and it provided better QoS parameters. It improved the service delay and throughput compared to other approaches.

REFERENCES:

- [1] D. Wu, Y.T. Hou, W.Z.Y.-Q. Zhang, and J.M. Peha, "Streaming Video over the Internet: Approaches and Directions," IEEE Trans. Circuits and Systems for Video Technology, vol. 11, no. 3, pp. 282-300, Mar. 2001.
- [2] S.-Y. Choi and A.B. Whinston, "The Future of e-Commerce: Integrate and Customize," Computer, vol. 32, no. 1, pp. 133-134, Jan. 1999.
- [3] J.W.S. Liu, W.-K. Shih, K.-J. Lin, R. Bettati, and J.-Y. Chung, "Imprecise Computations," Proc. IEEE, vol. 82, Jan. 1994.
- [4] S. Zilberstein and S.J. Russell, "Optimal Composition of Real-Time Systems," Artificial Intelligence, vol. 82, nos. 1-2, pp. 181-213, 1996.
- [5] X. Chen and A.M. K. Cheng, "An Imprecise Algorithm for Real-Time Compressed Image and Video Transmission," Proc. Int'l Conf. Computer Comm. and Networks (ICCCN), 1997.
- [6] M. Yannakakis, "Perspectives on Database Theory," Proc. Ann. Symp. Foundations of Computer Science, 1995.
- [7] T.F. Abdelzaher, K.G. Shin, and N. Bhatti, "Performance Guarantees for Web Server End-Systems: A Control-Theoretical Approach," IEEE Trans. Parallel and Distributed Systems, vol. 13, no. 1, pp. 80-96, Jan. 2002.
- [8] M. Arlitt and T. Jin, "A Workload Characterization Study of the 1998 World Cup Web Site," IEEE Network, vol. 14, no. 3, pp. 30-37, May-June 2000.
- [9] N. Bhatti and R. Friedrich, "Web Server Support for Tiered Services," IEEE Network, vol. 13, no. 5, pp. 64-71, 1999.
- [10] P. Bhoj, S. Ramanathan, and S. Singhal, "Web2K: Bringing QoS to Web Servers," Technical Report HPL-2000-61, HP Laboratories, May 2000.
- [11] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," IETF, RFC 2475, Dec. 1998.
- [12] J.M. Blanquer, A. Batchelli, K. Schausser, and R. Wolski, "Quorum: Flexible Quality of Service for Internet Services," Proc. Symp. Networked Systems Design and Implementation, 2005.
- [13] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: An Overview," RFC 1633, June 1994.
- [14] J. Almeida, M. Dabu, A. Manikutty, and P. Cao, "Providing Differentiated Levels of Service in Web Content Hosting," Proc. ACM SIGMETRICS Workshop Internet Server Performance, pp. 91-102, 1998.
- [15] C. Dovrolis, D. Stiliadis, and P. Ramanathan, "Proportional Differentiated Services: Delay Differentiation and Packet Scheduling," IEEE/ACM Trans. Networking, vol. 10, no. 1, pp. 12-26, 2002.
- [16] L. Eggert and J. Heidemann, "Application-Level Differentiated Services for Web Servers," World Wide Web J., vol. 2, no. 3, pp. 133-142, 1999.
- [17] A. Kamra, V. Misra, and E. Nahum, "Yaksha: A Self Tuning Controller for Managing the Performance of 3-Tiered Websites," Proc. Int'l Workshop Quality of Service, pp. 47-56, 2004.
- [18] M. Karlsson, C. Karamanolis, and X. Zhu, "Triage: Performance Isolation and Differentiation for Storage Systems," Proc. Int'l Workshop Quality of Service, pp. 67-76, 2004.
- [19] B. Li and K. Nahrstedt, "A Control-Based Middleware Framework for Quality of Service Adaptations," IEEE J. Selected Areas in Comm., vol. 17, no. 9, pp. 1632-1650, Sept. 1999.
- [20] Y. Lu, T.F. Abdelzaher, C. Lu, L. Sha, and X. Liu, "Feedback Control with Queueing-Theoretic Prediction for Relative Delay Guarantees in Web Servers," Proc. IEEE Real-Time and Embedded Technology and Applications Symp., pp. 208-217, May 2003.
- [21] P. Pradhan, R. Tewari, S. Sahu, A. Chandra, and P. Shenoy, "An Observation-Based Approach towards Self-Managing Web Servers," Proc. Int'l Workshop Quality of Service, 2002.
- [22] L. Sha, X. Liu, Y. Lu, T.F. Abdelzaher, "Queueing Model Based Network Server Performance Control," Proc. IEEE Real-Time Systems Symp., pp. 81-90, 2002.
- [23] V. Sundaram and P. Shenoy, "A Practical Learning-Based Approach for Dynamic Storage Bandwidth Allocation," Proc. Int'l Workshop Quality of Service, 2003.
- [24] B. Urgaonkar, P. Shenoy, "Cataclysm: Handling Extreme Overloads in Internet Applications," Proc. Int'l World Wide Web Conf., May 2005.
- [25] K.D. Kang, J. Oh, and S.H. Son, "Chronos: Feedback Control of a Real Database System Performance," IEEE Real-Time Systems Symp., 2007.
- [26] "Oracle Berkeley DB Product Family," <http://www.oracle.com/database/berkeleydb/index.html>. 2010.
- [27] Jianbin Wei, and Cheng-Zhong Xu, "eQoS: Provisioning of Client-Perceived End-to-End QoS Guarantees in Web Servers", IEEE Transactions On Computers, Vol. 55, No. 12, December 2006.
- [28] Mehdi Amirijoo, Jorgen Hansson, and Sang Hyuk Son, "Specification and Management of QoS in Real-Time Databases Supporting Imprecise Computations", IEEE Transactions on Computers, Vol. 55, No. 3, March 2006.
- [29] Kyoung-Don Kang, Yan Zhou, and Jisu Oh, "Estimating and Enhancing Real-Time Data Service Delays: Control-Theoretic Approaches" IEEE Transactions On Knowledge And Data Engineering, Vol. 23, No. 4, April 2011.

K Direct Torque Control of Induction Motor Using Space Vector Modulation (SVM- DTC)

Hanumanji Kantari

Abstract: In this study, direct torque control (DTC) of induction motor is evaluated based on space vector modulation (SVM). DTC is a method to control machine with utilizing torque and flux of motor controlled. The torque and current ripple are occurred in the conventional DTC. Reason of undesired torque and current ripple is low number of voltage vectors applied to the motor controlled by the conventional DTC technique. SVM-DTC is a technique to reduce the ripple. SVM techniques have several advantages that are offering better DC bus utilization, lower torque ripple, lower total harmonic distortion in the AC motor current, lower switching loss, and easier to implement in the digital systems. Simulation results from the classical and improved DTC are presented and compared. Result shows that the torque, flux linkage and stator current ripple are decreased with the improved DTC.

I. INTRODUCTION

Induction Motors (IMs) are widely used in high performance drives. Its history is very extensive and also control is important in applications. There are many IMs in a number of industrial, commercial and domestic applications of variable speed drives. Since IMs demands well control performances: precise and quick torque and flux response, large torque at low speed, wide speed range, the drive control system is the most sensitive point of IMs [1].

DTC is method to control machine with utilizing torque and flux of motor controlled. The basic DTC scheme consists of two comparators having different features, switching table, Voltage Source Inverter (VSI), flux and torque estimation block and IM.

Like a every control method has some advantages and disadvantages, DTC method has too. Some of the advantages are lower parameters dependency, making the system more robust and easier implements and the disadvantages are difficult to control flux and torque at low speed, current and torque distortion during the change of the sector in d-q plane, variable switching frequency, a high sampling frequency needed for digital implementation of hysteresis controllers, high torque ripple. The torque ripple generates noise and vibrations, causes errors in sensor less motor drives, and associated current ripples are in turn responsible for the EMI. The reason of the high current and torque ripple in DTC is the presence of hysteresis comparators together the limited number of available voltage vectors. To reduce the torque ripple, the stator flux vector change, which is demanded to compensate the torque and flux errors, determination should be done and also any voltage vector should be produced by the control mechanism. If a higher number of voltage vectors than those used in conventional DTC is used, the favorable motor control can be obtained. Because of complexity of power and control circuit, this approach is not satisfactory for low or medium power applications. One of the methods to increase the number of available vectors is an on-line modulation between active and null vectors.

In order to overcome the problem, the SVM-DTC and DSVM-DTC methodologies were proposed. The basis of the SVM-DTC methodology is the calculation of the required voltage space vector to compensate the flux and torque errors exactly by using a predictive technique and then its generation using the SVM at each sample period. There are different SVM techniques, which are Direct-Reverse SVM, Direct- Direct SVM, Direct-Direct with $V_{null}=[000]$, Direct-Direct with $V_{null}=[111]$ in industry applications. The choice of the SVM technique to be used will depend on the optimization criteria under consideration, whether it is the torque/current ripple, the harmonic losses or the switching losses.

II. INDUCTION MOTOR

An induction motor (IM) is a type of asynchronous AC motor where power is supplied to the rotating device by means of electromagnetic induction. Other commonly used name is squirrel cage motor due to the fact that the rotor bars with short circuit rings resemble a squirrel cage (hamster wheel).

An electric motor converts electrical power to mechanical power in its rotor (rotating part). There are several ways to supply power to the rotor. In a DC motor this power is supplied to the armature directly from a DC source, while in an induction motor this power is induced in the rotating device. An induction motor is sometimes called a rotating transformer because the stator (stationary part) is essentially the primary side of the transformer and the rotor (rotating part) is the secondary side. Induction motors are widely used, especially polyphase induction motors, which are frequently used in industrial drives.

Induction motors are now the preferred choice for industrial motors due to their rugged construction, absence of brushes (which are required in most DC motors) and — thanks to modern power electronics — the ability to control the speed of the motor.

Principle of operation and comparison to synchronous motors

A 3-phase power supply provides a rotating magnetic field in an induction motor.

The basic difference between an induction motor and a synchronous AC motor is that in the latter a current is supplied onto the rotor. This then creates a magnetic field which, through magnetic interaction, links to the rotating magnetic field in the

stator which in turn causes the rotor to turn. It is called synchronous because at steady state the speed of the rotor is the same as the speed of the rotating magnetic field in the stator.

By way of contrast, the induction motor does not have any direct supply onto the rotor; instead, a secondary current is induced in the rotor. To achieve this, stator windings are arranged around the rotor so that when energized with a polyphase supply they create a rotating magnetic field pattern which sweeps past the rotor. This changing magnetic field pattern can induce currents in the rotor conductors. These currents interact with the rotating magnetic field created by the stator and the rotor will turn.

However, for these currents to be induced, the speed of the physical rotor and the speed of the rotating magnetic field in the stator must be different, or else the magnetic field will not be moving relative to the rotor conductors and no currents will be induced. If by some chance this happens, the rotor typically slows slightly until a current is re-induced and then the rotor continues as before. This difference between the speed of the rotor and speed of the rotating magnetic field in the stator is called slip. It is unit less and is the ratio between the relative speeds of the magnetic field as seen by the rotor (the slip speed) to the speed of the rotating field. Due to this an induction motor is sometimes referred to as an asynchronous machine.

III. Construction

The stator consists of wound 'poles' that carry the supply current to induce a magnetic field that penetrates the rotor. In a very simple motor, there would be a single projecting piece of the stator (a salient pole) for each pole, with windings around it; in fact, to optimize the distribution of the magnetic field, the windings are distributed in many slots located around the stator, but the magnetic field still has the same number of north-south alternations. The number of 'poles' can vary between motor types but the poles are always in pairs (i.e. 2, 4, 6, etc.).

Induction motors can be built to run on either single-phase or three-phase power. Single-phase power is more widely available in residential buildings, but cannot produce a rotating field in the motor (the field merely oscillates back and forth), so single-phase induction motors must incorporate some kind of starting mechanism to produce a rotating field. They would, using the simplified analogy of salient poles, have one salient pole per pole number; a four-pole motor would have four salient poles. Three-phase motors have three salient poles per pole number, so a four-pole motor would have twelve salient poles. This allows the motor to produce a rotating field, allowing the motor to start with no extra equipment and run more efficiently than a similar single-phase motor.

There are two types of rotor:

- **Squirrel-cage rotor**

The most common rotor is a squirrel-cage rotor. It is made up of bars of either solid copper (most common) or aluminum that span the length of the rotor, and are connected through a ring at each end. The rotor bars in squirrel-cage induction motors are not straight, but have some skew to reduce noise and harmonics.

Slip ring rotor

A slip ring rotor replaces the bars of the squirrel-cage rotor with windings that are connected to slip rings. When these slip rings are shorted, the rotor behaves similarly to a squirrel-cage rotor; they can also be connected to resistors to produce a high-resistance rotor circuit, which can be beneficial in starting.

Speed control

The rotational speed of the rotor is controlled by the number of pole pairs (number of windings in the stator) and by the frequency of the supply voltage. Before the development of cheap power electronics, it was difficult to vary the frequency to the motor and therefore the uses for the induction motor were limited.

The general term for a power electronic device that controls the speed as well as other parameters is *inverter*. A typical unit will take the mains AC supply, rectify and smooth it into a "link" DC voltage, and, then convert it into the desired AC waveform.

Because the induction motor has no brushes and is easy to control, many older DC motors are being replaced with induction motors and accompanying inverters in industrial applications.

Starting of induction motors

Three Phase

Direct-on-line (DOL) starting

The simplest way to start a three-phase induction motor is to connect its terminals to the line. In an induction motor, the magnitude of the induced EMF in the rotor circuit is proportional to the stator field and the slip speed (the difference between synchronous and rotor speeds) of the motor, and the rotor current depends on this EMF. When the motor is started, the slip speed is equal to the synchronous speed, as the rotor speed is zero (slip equal to 1), so the induced EMF in the rotor is large. As a result, a very high current flows through the rotor. This is similar to a transformer with the secondary coil short circuited, which causes the primary coil to draw a high current from the mains. When an induction motor starts DOL, a very high current is drawn by the stator, on the order of 5 to 9 times the full load current. This high current can, in some motors, damage the windings; in addition, because it causes heavy line voltage drop, other appliances connected to the same line may be affected by the voltage fluctuation. To avoid such effects, several other strategies are employed for starting motors.

Star-delta starters

An induction motor's windings can be connected to a 3-phase AC line in two different ways: *wye* (*star* in Europe), where the windings are connected from phases of the supply to the neutral, and *delta* (sometimes *mesh* in Europe), where the windings are connected between phases of the supply. A delta connection results in a higher voltage to the windings than a wye connection (the voltage is multiplied by $\sqrt{3}$). A star-delta starter initially connects the motor in wye, which produces a lower starting current than delta, then switches to delta when the motor has reached a set speed. Their disadvantages over DOL starting are:

- Lower starting torque, which may be a serious issue with pumps or any devices with significant breakaway torque
- Increased complexity: more contactors and some sort of speed switch are needed
- Two shocks to the motor (one for the initial start and another when the motor switches from wye to delta)

Variable-frequency drives

VFDs can be of considerable use in starting as well as running motors. A VFD can easily start a motor at a lower frequency than the AC line, as well as a lower voltage, so that the motor starts with full rated torque and with no inrush of current. The rotor circuit's impedance increases with slip frequency, which is equal to supply frequency for a stationary rotor, so running at a lower frequency actually increases torque.

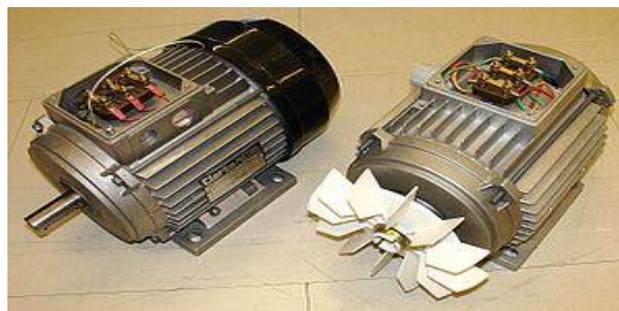
Single Phase

In a single phase induction motor, it is necessary to provide a starting circuit to start rotation of the rotor. If this is not done, rotation may be commenced by manually giving a slight turn to the rotor. The single phase induction motor may rotate in either direction and it is only the starting circuit which determines rotational direction.

For small motors of a few watts it is usual to "shade" the stator poles by means of a single turn of heavy copper wire around one corner of the pole. The current induced in the single turn is out of phase with the supply current and so causes an out-of-phase component in the magnetic field, which imparts to the field sufficient rotational character to start the motor. Starting torque is very low and efficiency is also reduced. Such shaded-pole motors are typically used in low-power applications with low or zero starting torque requirements, such as desk fans and record players.

Larger motors are provided with a second stator winding which is fed with an out-of-phase current to create a rotating magnetic field. The out-of-phase current may be derived by feeding the winding through a capacitor, or it may derive from the winding having different values of inductance and resistance from the main winding.

In some designs the second winding is disconnected once the motor is up to speed, usually either by means of a switch operated by centrifugal force acting on weights on the motor shaft, or by a positive temperature coefficient thermistor which after a few seconds of operation heats up and increases its resistance to a high value, reducing the current through the second winding to an insignificant level. Other designs keep the second winding continuously energized during running, which improves torque.



Three-phase induction motors



A squirrel-cage AC motor

Applications

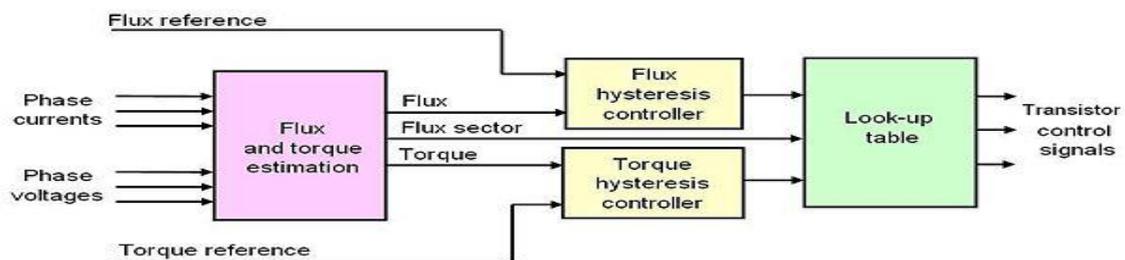
A wide variety of induction motors are available and are currently in use throughout a range of industrial applications. Induction motors typically include a rotor that rotates in response to a rotating magnetic flux generated by alternating current (AC) in a stator associated with the rotor. A rotational speed differential between the rotor and the rotating flux induces a current through a rotor cage. A rotor cage consists of a single aluminum casting having several conductive bars that run axially through the rotor and are joined at each end by two conductive end rings. Current induced in the bars creates a magnetic flux that opposes that of the stator, thus providing the rotor with rotational torque. The stator and rotor may be mechanically and electrically configured in a variety of manners depending upon a number of factors, including: the application, the power available to drive the motor, and so forth. The induction motors are divided again into an inner rotor type induction motor and an outer rotor type induction motor in accordance with relative positions of rotors and stators. The inner rotor type induction motor is generally applied to a washing machine or the like, and includes the rotor inside the stator. The ability to accurately determine the speed of a rotating rotor with respect to a stationary stator within an induction motor is vitally important to the every day operations of induction motors. A flat induction motor is a motor which has a disk-shaped stator and rotor placed coaxially around a rotating shaft with their surfaces opposing each other. Normally, the stator core and rotor core each have a spiral winding structure made of a magnetic steel strip. A plurality of open slots is formed in the winding structures from the outer edge toward the rotating shaft at equal intervals, leaving part of the magnetic steel strips. Single phase induction motors are normally provided with a cage type rotor and a coiled stator having two windings, one being for the running coil and the other for the starting coil. Single phase induction motors are widely used, due to their simplicity, strength and high performance. They are used in household appliances, such as refrigerators, air conditioners, hermetic compressors, washing machines, pumps, fans, as well as in some industrial applications. Linear induction motors are widely used in a number of industries and present certain advantages over rotary motors, particularly where propulsion along a predetermined path or guide way is required

IV. Direct Torque Control

Direct torque control (DTC) is one method used in variable frequency drives to control the torque (and thus finally the speed) of three-phase AC electric motors. This involves calculating an estimate of the motor's magnetic flux and torque based on the measured voltage and current of the motor.

Method

Stator flux linkage is estimated by integrating the stator voltages. Torque is estimated as a cross product of estimated stator flux linkage vector and measured motor current vector. The estimated flux magnitude and torque are then compared with their reference values. If either the estimated flux or torque deviates from the reference more than allowed tolerance, the transistors of the variable frequency drive are turned off and on in such a way that the flux and torque will return in their tolerance bands as fast as possible. Thus direct torque control is one form of the hysteresis or bang-bang control.



This control method implies the following properties of the control:

- Torque and flux can be changed very fast by changing the references
- The step response has no overshoot
- No coordinate transforms are needed, all calculations are done in stationary coordinate system
- No separate modulator is needed, the hysteresis control defines the switch control signals directly
- There are no PI current controllers. Thus no tuning of the control is required
- The switching frequency of the transistors is not constant. However, by controlling the width of the tolerance bands the average switching frequency can be kept roughly at its reference value. This also keeps the current and torque ripple small. Thus the torque and current ripple are of the same magnitude than with vector controlled drives with the same switching frequency.
- Due to the hysteresis control the switching process is random by nature. Thus there are no peaks in the current spectrum. This further means that the audible noise of the machine is low
- The intermediate DC circuit's voltage variation is automatically taken into account in the algorithm (in voltage integration). Thus no problems exist due to dc voltage ripple (aliasing) or dc voltage transients
- Synchronization to rotating machine is straightforward due to the fast control; Just make the torque reference zero and start the inverter. The flux will be identified by the first current pulse

- Digital control equipment has to be very fast in order to be able to prevent the flux and torque from deviating far from the tolerance bands. Typically the control algorithm has to be performed with 10 - 30 microseconds or shorter intervals. However, the amount of calculations required is small due to the simplicity of the algorithm
- The current and voltage measuring devices have to be high quality ones without noise and low-pass filtering, because noise and slow response ruins the hysteresis control
- In higher speeds the method is not sensitive to any motor parameters. However, at low speeds the error in stator resistance used in stator flux estimation becomes critical

The direct torque method performs very well even without speed sensors. However, the flux estimation is usually based on the integration of the motor phase voltages. Due to the inevitable errors in the voltage measurement and stator resistance estimate the integrals tend to become erroneous at low speed. Thus it is not possible to control the motor if the output frequency of the variable frequency drive is zero. However, by careful design of the control system it is possible to have the minimum frequency in the range 0.5 Hz to 1 Hz that is enough to make possible to start an induction motor with full torque from a standstill situation. A reversal of the rotation direction is possible too if the speed is passing through the zero range rapidly enough to prevent excessive flux estimate deviation. If continuous operation at low speeds including zero frequency operation is required, a speed or position sensor can be added to the DTC system. With the sensor high accuracy of the torque and speed control can be maintained in the whole speed range.

DIRECT TORQUE CONTROL FOR INDUCTION MOTOR

MAIN FEATURES OF DTC

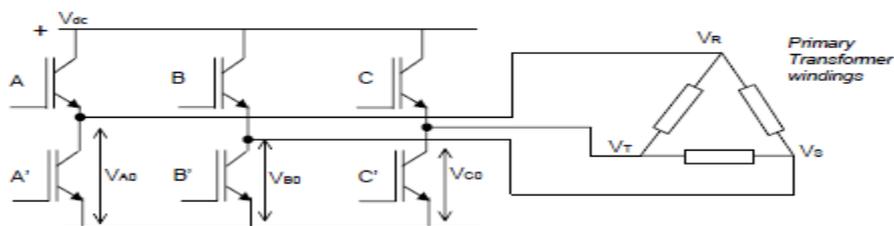
- Decoupled control of torque and flux
- Absence of mechanical transducers
- Current regulator, PWM pulse generation, PI control of flux and torque and co-ordinate transformation is not required
- Very simple control scheme and low computational time
- Reduced parameter sensitivity

V. SPACE VECTOR MODULATION

PWM, as discussed in paragraph 4, can be generated by analogue or digital control electronics. The advantages of digital controls over analogue are:

- Stability (no drift, offsets or aging effects)
- Precision (noise immunity)
- Flexibility (can be customized by changing software)

Even if done digitally, significant computing time is required, as the PWM signals have to be calculated in real time. By using Space Vector Modulation this calculation process is simplified. As it is simplified, less computing time is required, and therefore better performance can be obtained.



As we have 3 switches (A, B and C) there are 8 different switching combinations. For example A on, B off, C on. We can analyse for all switching combinations the phase-phase voltages on the primary winding of the transformer. If we look to the voltages that need to be generated between each leg (fig 8) we can split that into 8 different ‘spaces’, S1, S2 etc. For each space we can use the table to define which switching states the inverter must use to obtain those voltages. For example for period (‘space’) S1, we see that the VRS is positive, VST is negative and VTR is positive. Besides the fact that these voltages are positive or negative they also reach a point where they are zero. From the table it can be read that in period S1 the inverter needs to be switching between state Nr. 1, 2, 8 and 6.

Nr	A	B	C	V _{A0}	V _{B0}	V _{C0}	V _{RS}	V _{ST}	V _{TR}	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆
1	0	0	0	0	0	0	0	0	0	X	X	X	X	X	X
2	0	0	1	0	0	V _{dc}	0	-V _{dc}	V _{dc}	X					X
3	0	1	0	0	V _{dc}	0	-V _{dc}	V _{dc}	0				X	X	
4	0	1	1	0	V _{dc}	V _{dc}	-V _{dc}	0	V _{dc}					X	X
5	1	0	0	V _{dc}	0	0	V _{dc}	0	-V _{dc}		X	X			
6	1	0	1	V _{dc}	0	V _{dc}	V _{dc}	-V _{dc}	0	X	X				
7	1	1	0	V _{dc}	V _{dc}	0	0	V _{dc}	-V _{dc}			X	X		
8	1	1	1	V _{dc}	V _{dc}	V _{dc}	0	0	0	X	X	X	X	X	X

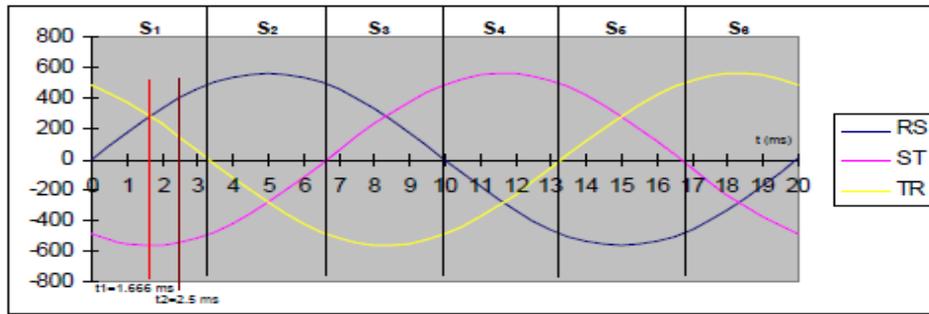
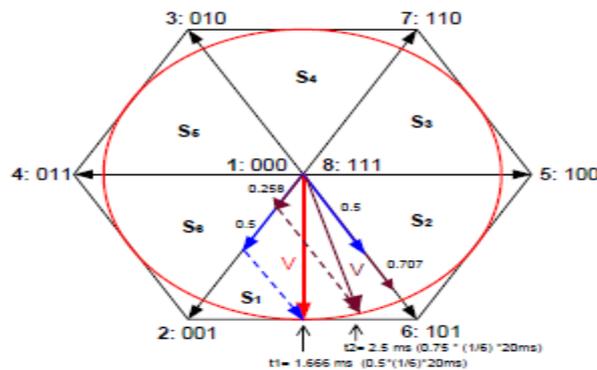


Fig above in combination with the table can also be represented in a hexagon, as shown in fig



In period S1 the inverter will switch between switching states 1,2, 8 and 6, in the period S2 between: 1, 6, 8 and 5. etc. One full period (in case of 50Hz 1 period is 20ms) can be represented by a vector V making one full circle in the hexagon. The period showed in figure 8 starts in the hexagon at state 2 and is then turning counter clockwise. As shown in fig 9 at t1=1.666 ms the vector V (red) is built up by staying 25% of the time in state 1, 25% of the time in state 2, 25% of the time in state 8 and 25% of the time in state 6. This switching pattern in one period of the switching frequency is shown in figure 10. For t2= 2.5 ms the required time per switching state is calculated as follows:

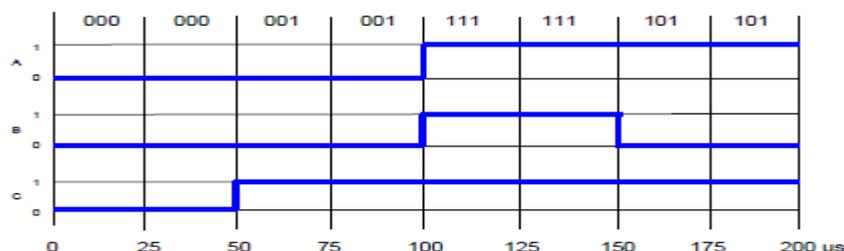
$$V_{RS} = \sin(2\pi \cdot f \cdot t) \quad \text{at } t=2.5\text{ms} \quad V_{RS} = 0.707 \quad (73.2\% \text{ of } 0.965 \text{ which is the total})$$

$$V_{TR} = \sin(2\pi \cdot f \cdot t - 4/3\pi) \quad \text{at } t=2.5\text{ms} \quad V_{TR} = 0.258 \quad (26.8\% \text{ of } 0.965 \text{ which is the total})$$

Please note that V_{RS} and V_{TR} are normalized to 1 where 1 corresponds with 665V, the maximum value ($400\sqrt{2}$)

So within one switching period the inverter needs to stay

- 25.0% of the time on state 1 (always)
- 26.8% * 50% = 13.4% of the time on state 2
- 25.0% of the time on state 8 (always)
- 73.2% * 50% = 36.6% of the time on state 6



shows the switching states as shown at t1 in fig 9 and visualizes the switching states of leg A,B and C during 1 period. (1 period is 200 microseconds, as the switching frequency is 5kHz).

In a UPS Space Vector Modulation can be implemented by means of DSPs (Digital Signal Processors). These are microprocessors which are suitable to process analogue signals. DSPs provide fast sampling rates which make them ideally suited to UPS control.

By using an isolation transformer on the output of the inverter, special modulation techniques can be applied. The modulation index can be increased with 15% by adding a 3rd harmonic (150Hz in case of 50Hz output) to the signal on each leg of the transformer. As the transformer is in dy configuration, this 3rd harmonic will not appear on the output voltage. This optimisation is not possible in transformerless inverters (UPS).

BENEFITS

Space Vector Modulation for a three phase UPS inverter makes it possible to adapt the switching behaviour to different situations such as: half load, full load, linear load, non-linear load, static load, pulsating load, etc. In combination with a zig-zag three phase transformer in the output this provides the following advantages:

- Very low values can be reached for the output voltage THD (<2% for linear loads., <3% for non linear loads)
- Robust dynamic response (<3% deviation at 100% load step, recovery time to <1%: <20ms)
- The efficiency of the inverter can be optimized, for each load condition.
- Because of the strong regulation in combination with a zig-zag transformer the inverter can accept a 100% unbalanced load and maintain the performance
- VM enables more efficient use of the DC voltage (15% more than conventional PWM techniques, so the inverter will accept a 15% lower DC voltage making full use of the available battery energy)
- By applying special modulation techniques the peak currents in the IGBTs can be reduced compared to similar inverters. This improves the MTBF of the inverter, since there is less thermal stress on the IGBT chip.
- By changing the switching behaviour of the inverter, the audible noise can also be influenced and therefore be minimized.

Space Vector Modulation provides excellent output performance, optimized efficiency, and high reliability compared to similar inverters with conventional Pulse Width Modulation.

VI. INSULATED GATE BIPOLAR TRANSISTOR (IGBT)

IGBT has been developed by combining into it the best qualities of both BJT and PMOSFET. Thus an IGBT possesses high input impedance like a PMOSFET and has low on-state power loss as in a BJT. Further, IGBT is free from second breakdown problem present in BJT. All these merits have made IGBT very popular amongst power-electronics engineers. IGBT is also known as metal oxide insulated gate transistor (MOSIGT), conductively-modulated field effect transistor (COMFET) or gain-modulated FET(GEMFET). It was also initially called insulated gate transistor (IGT).

The **insulated-gate bipolar transistor** or **IGBT** is a three-terminal power semiconductor device, noted for high efficiency and fast switching. It switches electric power in many modern appliances: electric cars, variable speed refrigerators, air-conditioners, and even stereo systems with digital amplifiers. Since it is designed to rapidly turn on and off, amplifiers that use it often synthesize complex waveforms with pulse width modulation and low-pass filters.

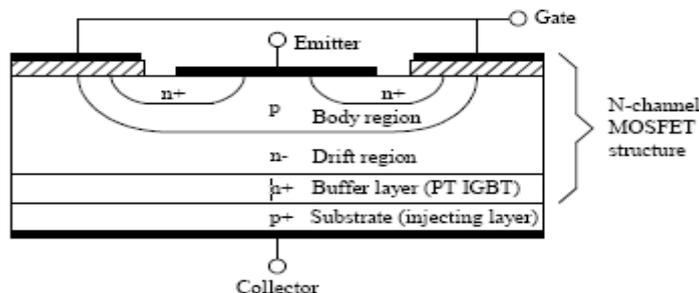
The IGBT combines the simple gate-drive characteristics of the MOSFETs with the high-current and low-saturation-voltage capability of bipolar transistors by combining an isolated-gate FET for the control input, and a bipolar power transistor as a switch, in a single device. The IGBT is used in medium- to high-power applications such as switched-mode power supply, traction motor control and induction heating. Large IGBT modules typically consist of many devices in parallel and can have very high current handling capabilities in the order of hundreds of amps with blocking voltages of 6,000 V.

The IGBT is a fairly recent invention. The first-generation devices of the 1980s and early 1990s were relatively slow in switching, and prone to failure through such modes as latch up and secondary breakdown. Second-generation devices were much improved, and the current third-generation ones are even better, with speed rivaling MOSFETs, and excellent ruggedness and tolerance of over loads.

Basic Structure

Fig illustrates the basic structure of an IGBT. It is constructed virtually in the same manner as a power MOSFET. There is, however, a major difference in the substrate. The n+ layer substrate at the drain in a PMOSFET is now substituted in the IGBT by a p+ layer substrate called collector C. Like a power MOSFET, an IGBT has also thousands of basic structure cell connected approximately on a single chip of silicon.

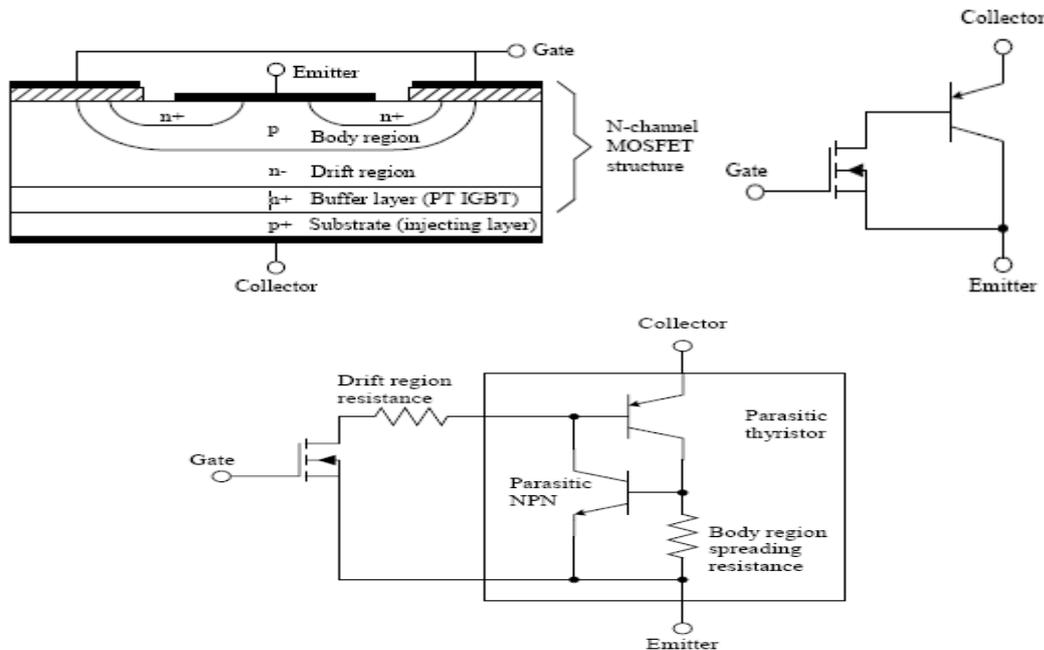
In IGBT, p+ substrate is called injection layer because it injects holes into n- layer. The n- layer is called drift region. As in other semiconductor devices, thickness of n- layer determines the voltage blocking capability of IGBT. The p layer is called body of IGBT. The n- layer in between p+ and p regions serves to accommodate the depletion layer of pn-junction, i.e. junction J2.



N-Channel IGBT Cross Section

Equivalent Circuit

An examination of reveals that if we move vertically up from collector to emitter. We come across p+, n-, p layer s. Thus, IGBT can be thought of as the combination of MOSFET and p+ n- p layer s. Thus, IGBT can be thought of as the combination of MOSFET and p+ n- p transistor Q1 .Here Rd is resistance offered by n- drift region. Approximate equivalent circuit of an IGBT.



Exact equivalent circuit

The existence of another path from collector to emitter, this path is collector, p+ n-, p (n-channel), n+ and emitter. There is, thus, another inherent transistor Q2 as n- pn+ in the structure of IGBT. The interconnection between two transistors Q1 and Q2.This gives the complete equivalent circuit of an IGBT. Here R_{by} is the existence offered by p region to flow of hole current I_h.

The two transistor equivalent circuit illustrates that an IGBT structure has a parasitic thyristor in it. Parasitic thyristor is shown in line.

Working

When collector is made positive with respect to emitter, IGBT gets forward biased. With no voltage between gate and emitter, two junctions between n- region and p region (i.e. junction J2) are reversed biased; so no current flows from collector to emitter

When gate is made positive with respect to emitter by voltage V_G, with gate-emitter voltage more than the threshold voltage V_{GET} of IGBT, an n-channel or inversion layer, is formed in the upper part of p region just beneath the gate, as in PMOSFET. This n- channel short circuits the n- region with n+ emitter regions. Electrons from the n+ emitter begin to flow to n- drift region through n-channel. As IGBT is forward biased with collector positive and emitter negative, p+ collector region injects holes into n- drift region .In short; n-drift region is flooded with electrons from p-body region and holes from p+ collector region. With this, the injection carrier density in n- drift region increases considerably and as a result, conductivity of n- region enhances significantly. Therefore, IGBT gets turned on and begins to conducts forward current I_C.

Current I_c, or I_e of two current components:

1. Holes current I_h due to injected holes flowing from collector,p+ n- p transistor Q1, p-body region resistance R_{by} and emitter .
2. Electronic current I_e due to injected electrons flowing from collector, or load, current I_{C=emitter} current I_e=I_h+I_e. Major component of collector current is electronic current I_e, i.e. main current path for collector, or load, current is through p+, n-, drift resistance R_d and n channel resistance R_{ch}. Therefore, the voltage drop in IGBT in its on-state is

$$V_{ce.on} = I_c \cdot R_{ch} + I_c \cdot R_d + V_{ji}$$

=voltage drop [in n- channel] + across drift in n- region + across forward biased p+ n- junction J1.

Here V_{ji} is usually 0.7 to 1v as in a p-n diode. The voltage drop I_c. R_{ch} is due to n-channel resistance, almost the same as in a PMOSFET. The voltage drop V_{df} = I_c.R_d in UGBT is much less than that in PMOSFET. It is due to substantial increase in the conductivity caused by injection of electrons and holes in n- drift region. The conductivity increase is the main reason for the low on-state voltage drop in IGBT than that it is in PMOSFET.

Latch-up in IGBT

From the above that IGBT structure has two inherent transistors Q1 and Q2, which constitute a parasitic thyristor. When IGBT is on, the hole current flows through transistor p+ n- p and p- body resistance R_{by}. If load current I_c is large,

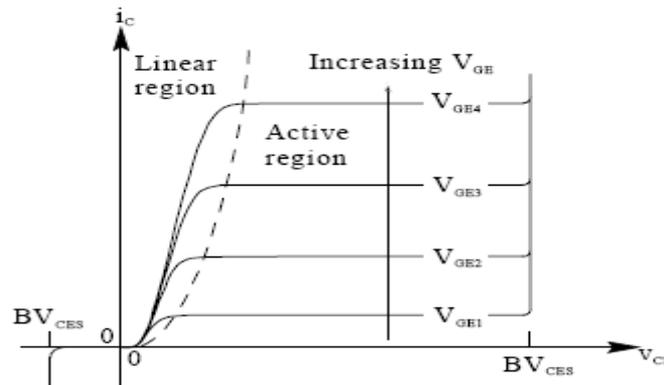
hole component of current I_h would also be large. This large current would increase the voltage drop $I_h \cdot R_{by}$ which may forward bias the base-p-emitter n+ junction of transistor Q2. As a consequence, parasitic transistor Q2 gets turned on which further facilitates in the turn-on of parasitic transistor p+ n- p labeled Q1. The parasitic thyristor, consisting of Q1 and Q2, eventually latches on through regenerative action, when sum of their current gains $\alpha_1 + \alpha_2$ reaches unity as in a conventional thyristor. With parasitic thyristor on, IGBT latches up and after this, collector emitter current is no longer under the control of gate terminal. The only way now to turn-off the latched up IGBT is by forced commutation of current as is done in a conventional thyristor. If this latch up is not aborted quickly, excessive power dissipation may destroy the IGBT. The latch up discussed here occurs when the collector current I_{ce} exceeds a certain critical value. The device manufacturers always specify the maximum permissible value of load current I_{ce} that IGBT can handle without latch up.

At present, several modifications in the fabrication techniques are listed in the literatures which are used to avoid latch-up in IGBTs. As such, latch up free IGBTs are available.

IGBT Characteristics

The circuit shows the various parameters pertaining to IGBT characteristics. Static I-V or output characteristics of an IGBT (n-channel type) show the plot of collector current I_c versus collector-emitter voltage V_{ce} for various values of gate-emitter voltages V_{GE1} , V_{GE2} etc. These characteristics are shown below. In the forward direction, the shape of the output characteristics is similar to that of BJT. But here the controlling parameter is gate-emitter voltage V_{GE} because IGBT is a voltage controlled device. When the device is off, junction J2 blocks forward voltage and in case reverse voltage appears across collector and emitter, junction J1 blocks it. V_{rm} is the maximum reverse breakdown voltage.

The transfer characteristic of an IGBT is a plot of collector current I_c versus gate-emitter voltage V_{GE} . This characteristic is identical to that of power MOSFET. When V_{GE} is less than the threshold voltage V_{GET} , IGBT is in the off state.



Static V-I characteristics

Switching Characteristics

Switching characteristics of an IGBT during turn-on and turn-off are sketched. The turn-on time is defined as the time between by instance of forward blocking to forward on-state. Turn-on time is composed of delay time t_{dn} and rise time t_r , i.e. $t_{on} = t_{dn} + t_r$. The delay time is defined as the time for the collector-emitter voltage to fall from V_{ce} to $0.9 V_{ce}$. Here V_{ce} is the initial collector-emitter voltage. Time t_{dn} may also be defined as the time for the collector current to rise from its initial leakage current I_{ce} to $0.1 I_c$. Here I_c is the final value of the collector current.

The rise time t_r is the time during which collector-emitter falls from $0.9 V_{CE}$ to $0.1 V_{CE}$. It is also defined as the time for the collector current to rise from $0.1 I_c$ to its final value I_c . After time t_{on} , the collector current I_c and the collector-emitter voltage falls to small value called conduction drop = V_{CES} where subscript s denotes saturated value.

The turn-off time is somewhat complex. It consists of three intervals

1. Delay time t_{df}
2. Initial fall time t_{f1}
3. Final time t_{f2}

i.e. $t_{off} = t_{df} + t_{f1} + t_{f2}$

The delay time is the time during which gate voltage falls from V_{GE} to threshold voltage V_{GET} . As V_{GE} falls to V_{GET} during t_{df} , the collector current falls from I_c to $0.9 I_c$. At the end of the t_{df} , collector-emitter voltage begins to rise. The first fall time T_{f1} is defined as the time during which collector current falls from 90 to 20 % of its initial value I_c , or the time during which collector-emitter voltage rises from V_{ce} to $0.1 V_{ce}$.

The final fall time t_{f2} is the time during which collector current falls from 20 to 10% of I_c , or the time during which collector-emitter voltage rises from $0.1 V_{CE}$ to final value V_{CE} .

Applications of IGBTs

IGBTs are widely used in medium power applications such as AC and DC motor drives, UPS systems, power supplies and drives for solenoids, relays and contactors. Though IGBTs are somewhat more expensive than BJTs, yet they are becoming popular because of lower gate-drive requirement, lower switching losses and smaller snubber circuit

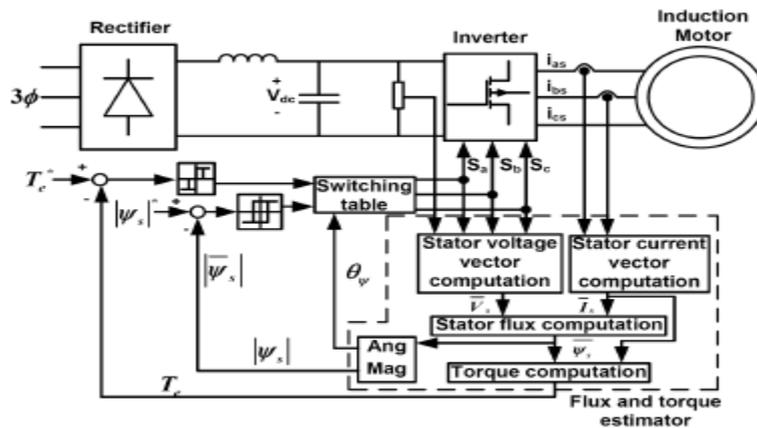
requirements. IGBT converter are more efficient with less size as well as cost, as compared to converters based on BJTs. Recently, IGBT inverter induction-motor drives using 15-20 KHZ. Switching frequency favour where audio-noise is objectionable. In most applications, IGBTs will eventually push out BJTs. At present, the state of the art IGBTs of 1200vots, 500 Amps ratings , 0.25-20 μ s turn off time with operating frequency are available.

VII. CONTROL METHOD

A) Conventional DTC

DTC method has been first proposed for induction machines. DTC technique introduced by Takahashi and Noguchi for low and medium power application and DTC technique introduced by Depenbrock for high power application are popular in industry. DTC strategy is quite different from that of the field orientation control (FOC) or vector control, which does not need complicated coordination transformations and decoupling calculation . The basic model of the conventional DTC induction motor scheme is shown in Figure 1. Two stator currents (i_{SA} and i_{SB}) and DC-bus voltage V_{DC} are sampled. d-q components of stator voltage and current space vectors in the stationary reference frame and also magnitude of the stator flux and electric torque are calculated as shown below

Basic DTC induction scheme



$$V_{zd} = \frac{2}{3} V_{dc} (S_A - \frac{S_B + S_C}{2})$$

$$V_{zq} = \frac{1}{\sqrt{3}} V_{dc} (S_A - S_C)$$

$$i_{zd} = i_{SA}$$

$$i_{zq} = \frac{i_{SA} + 2i_{SB}}{\sqrt{3}}$$

$$|\Psi_s| = \sqrt{\Psi_{zd}^2 + \Psi_{zq}^2}$$

$$T_e = \frac{3}{2} P (\Psi_{zd} i_{zq} - \Psi_{zq} i_{zd})$$

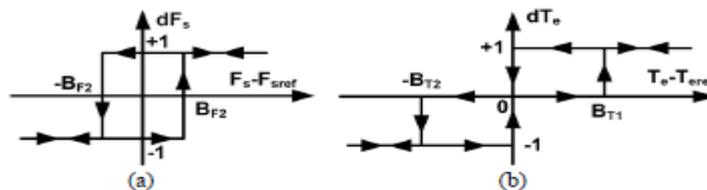
$$\Psi_{zd} = \int (V_{zd} - R_s i_{zd}) dt$$

$$\Psi_{zq} = \int (V_{zq} - R_s i_{zq}) dt$$

The magnitude of stator flux and electric torque calculated are compared with their reference values in the hysteresis comparators shown in Figure 2 and then the outputs of the comparators are fed to a switching table to select an appropriate inverter voltage vector. The switching table shown as Table I determine the voltage vector to apply based on the position of the stator flux and the required changes in stator flux magnitude and torque [8]. The selected voltage vector will be applied to the induction motor at the end of the sample time. In VSI, there are six equally spaced voltage vectors having the same amplitude and two zero voltage vectors. VSI voltage vectors are shown in Figure 3.

In DTC, torque and flux are controlled independently by selecting the optimum voltage space vector for entire switching period and the errors are maintained with in the hysteresis band [9]. In conventional DTC, only one vector is applied for the entire sampling period. So for small errors, the motor torque may exceed the upper\lower torque limit. Instead by using more than one vector with in the sampling period torque ripple can be reduced. The slip frequency can be controlled precisely by inserting zero vectors [10]. For the small the hysteresis band, frequency of operation of PWM inverter could be very high. The switching frequency always varies according to the width of hysteresis band.

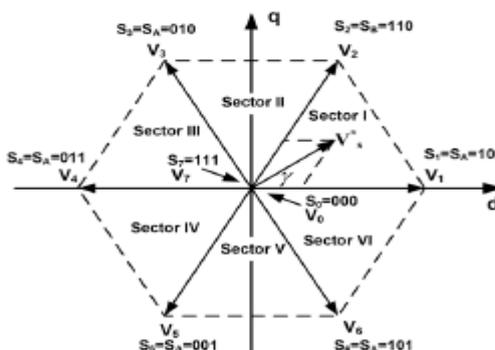
Hsteresis comparator a) stator flux b) torque



Switching table

Flux error position	Torque error position	Sector I	Sector II	Sector III	Sector IV	Sector V	Sector VI
1	1	V2(110)	V3(010)	V4(011)	V5(001)	V6(101)	V1(100)
	0	V7(111)	V0(000)	V7(111)	V0(000)	V7(111)	V0(000)
	-1	V6(101)	V1(100)	V2(110)	V3(010)	V4(011)	V5(001)
0	1	V3(010)	V4(011)	V5(001)	V6(101)	V1(100)	V2(110)
	0	V0(000)	V7(111)	V0(000)	V7(111)	V0(000)	V7(111)
	-1	V3(001)	V6(101)	V1(100)	V2(110)	V3(010)	V4(011)

VSI voltage vectors



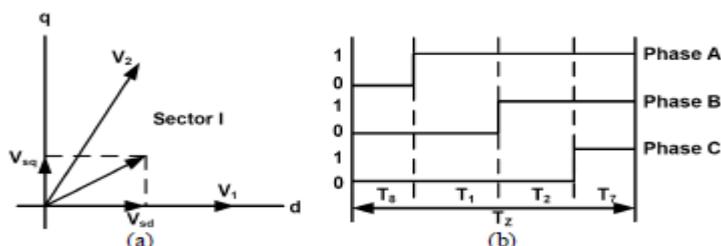
B) Space Vector Modulation

Common disadvantages of conventional DTC are high torque ripple and slow transient response to the step changes in torque during start-up. Several techniques have been developed to improve the torque performance. One of them is to reduce the ripples is based on SVM technique.

SVM was first presented by a group of German researched in the second half of the 1980s. Since then, a lot of work has been done on the theory and implementation of SVM techniques. SVM techniques have several advantages that are offering better DC bus utilization, lower torque ripple, lower

Total Harmonic Distortion (THD) in the AC motor current, lower switching losses, and easier to implement in the digital systems. At each cycle period, a preview technique is used to obtain the voltage space vector required to exactly compensate the flux and torque errors. The torque ripple for this SVM DTC is significantly improved and switching frequency is maintained constant [11]. SVM, based on the switching between two adjacent boundary active vectors and a zero vector during one switching period, T_z , and for a given reference voltage vector in the first sector (0-60°) is shown in Figure

SVM in the first sector (a) reference voltage vector (b) switching pattern for three phase modulation



The switching times can be calculated using the following equations

$$V_s \approx V_{sd} + jV_{sq}$$

where the vectors V_{sd} and V_{sq} are obtained from the appropriate voltage vectors for any given sector.

$$T_z \vec{V}_s = T_A \vec{V}_1 + T_B \vec{V}_2$$

$$T_A \sqrt{\frac{2}{3}} V_{dc} \begin{bmatrix} \cos 0^\circ \\ \sin 0^\circ \end{bmatrix} + T_B \sqrt{\frac{2}{3}} V_{dc} \begin{bmatrix} \cos(\pi/3) \\ \sin(\pi/3) \end{bmatrix}$$

$$= T_z \sqrt{\frac{2}{3}} V_{dc} a \begin{bmatrix} \cos \gamma \\ \sin \gamma \end{bmatrix}$$

$$T_A = T_z a \frac{\sin(\pi/3 - \gamma)}{\sin(\pi/3)}$$

$$T_B = T_z a \frac{\sin(\gamma)}{\sin(\pi/3)}$$

$$T_0 = T_7 = T_z - T_B - T_A$$

$$0 \leq \gamma \leq \pi/3, a = \frac{|\vec{V}_s|}{\sqrt{\frac{2}{3}} V_{dc}}$$

The reference vector $V^* s$, a constant magnitude and frequency in the steady-state, is sampled at equal time intervals of T_z . Within this sample time, the inverter is switched and made to remain at different switching states for different durations of time such that the average space vector generated within sample period is equal to the sampled value of the reference vector, both in terms of magnitude and angle. The switching states that can be used within T_z are the two zero states and the active states which are SA and SB, with vectors V_1 and V_2 respectively forming the start and the end boundaries of the sector as shown in Figure 4. The two switching states (SA and SB) are named active switching states. SA indicates the inverter switching states (001), (100), or (010) and SB indicates the inverter switching states (101), (110) or (011). Active vector times, T_A and T_B , are defined as the times due to the active switching states, SA and SB respectively. Null vector times T_0 and T_7 are defined as the times due to the null switching states S_0 , (000), and S_7 , (111) respectively. In DTC, with the space vector PWM technique, the DTC transient performance and robustness are preserved and the steady state torque ripple is reduced. Moreover, the inverter switching frequency is constant and totally controllable.

VIII. SIMULATION RESULTS

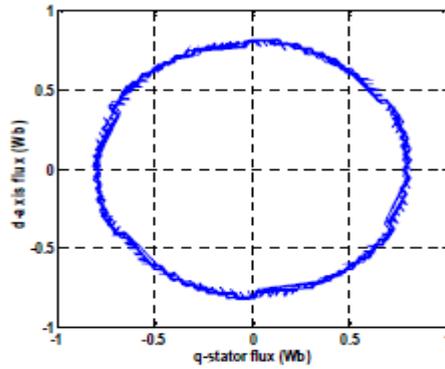
Two Matlab models were developed to examine the different control algorithm. One is used for the conventional DTC and the other for the modified DTC. The parameters of the induction motor are shown in Table II.

TABLE II. Parameters of the induction motor used

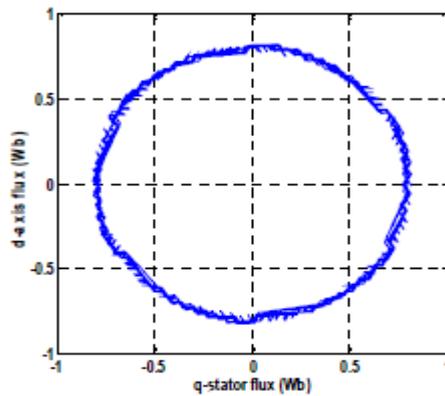
Stator resistance R_s	3 Ω
Rotor resistance R_r	5.6 Ω
Rotor Inductance L_r	0.0154 H
Mutual inductance M	0.0102 H
Stator Inductance L_s	0.0154 H
Phase voltage V	250 V
Base speed w_b	314 rad/s
Inertia J	0.002 kgm ²
Frequency f	50 Hz

The steady state behavior of induction motor with the conventional DTC and SVM-DTC are illustrated in Figure. Figure 5 and 6 show a comparison of stator flux space vector obtained using two control methods. As it is possible to see in Figure 7-12 an appreciable reduction of current, flux and torque ripple has been obtained using the SVM-DTC technique.

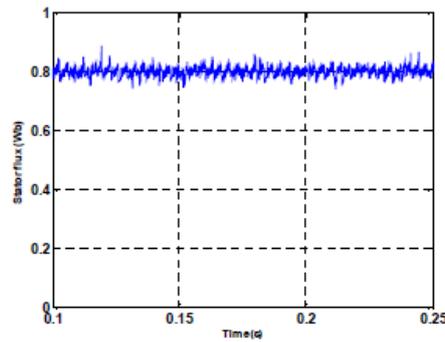
d-q staror flux with a conventional DTC scheme



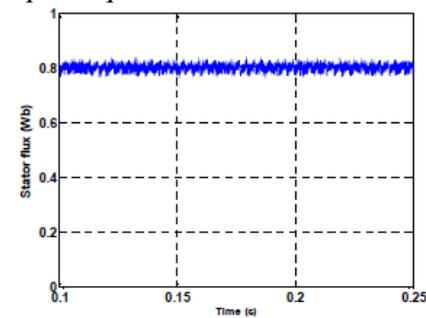
d-q stator flux with SVM-DTC scheme

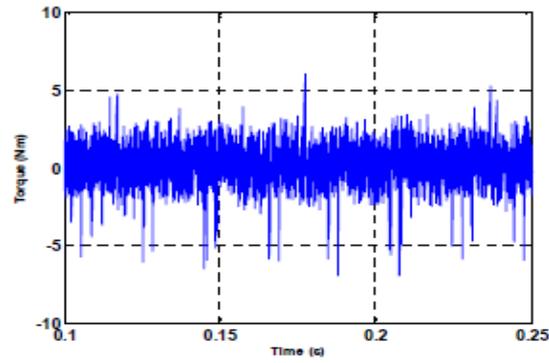


Stator flux magnitude with conventional DTC scheme

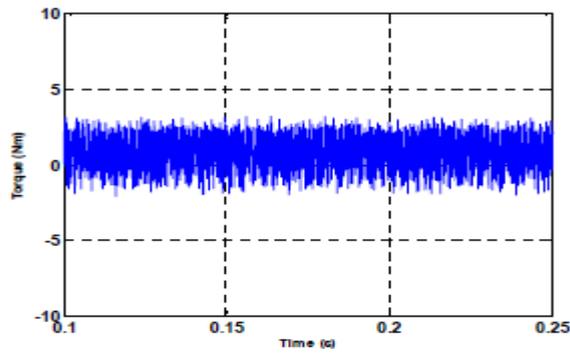


Out put torque with conventional DTC scheme

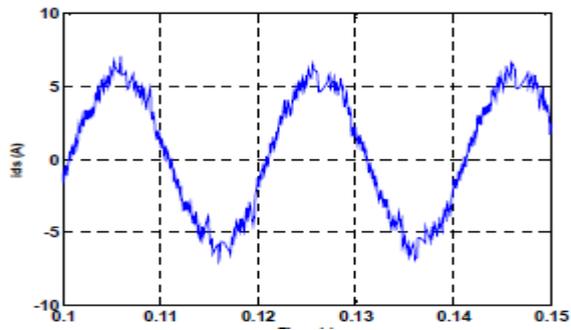




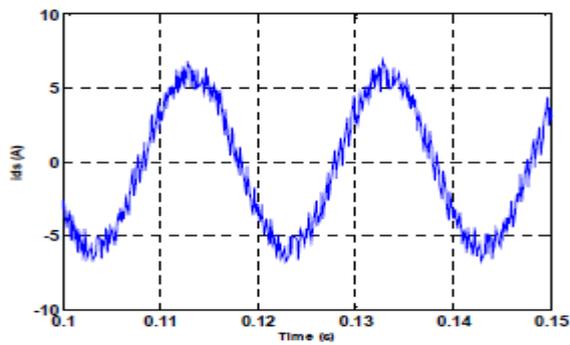
Out put torque with SVM-DTC scheme



Stator d-current with conventional DTC scheme

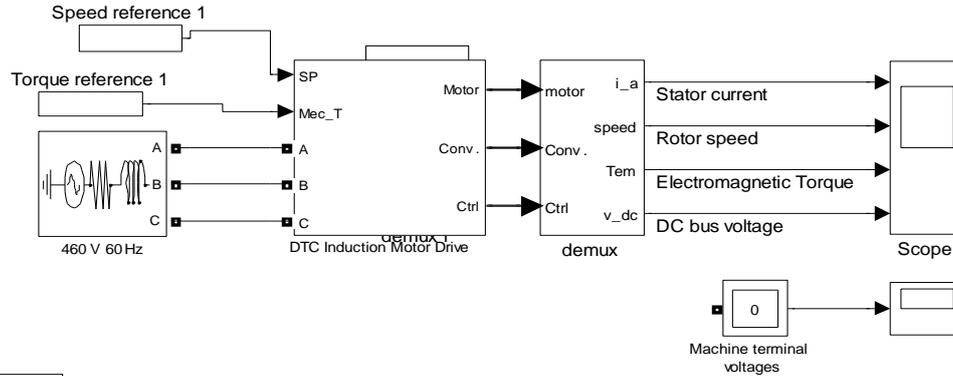


Stator d-current with SVM-DTC scheme



SIMULINK CIRCUIT:

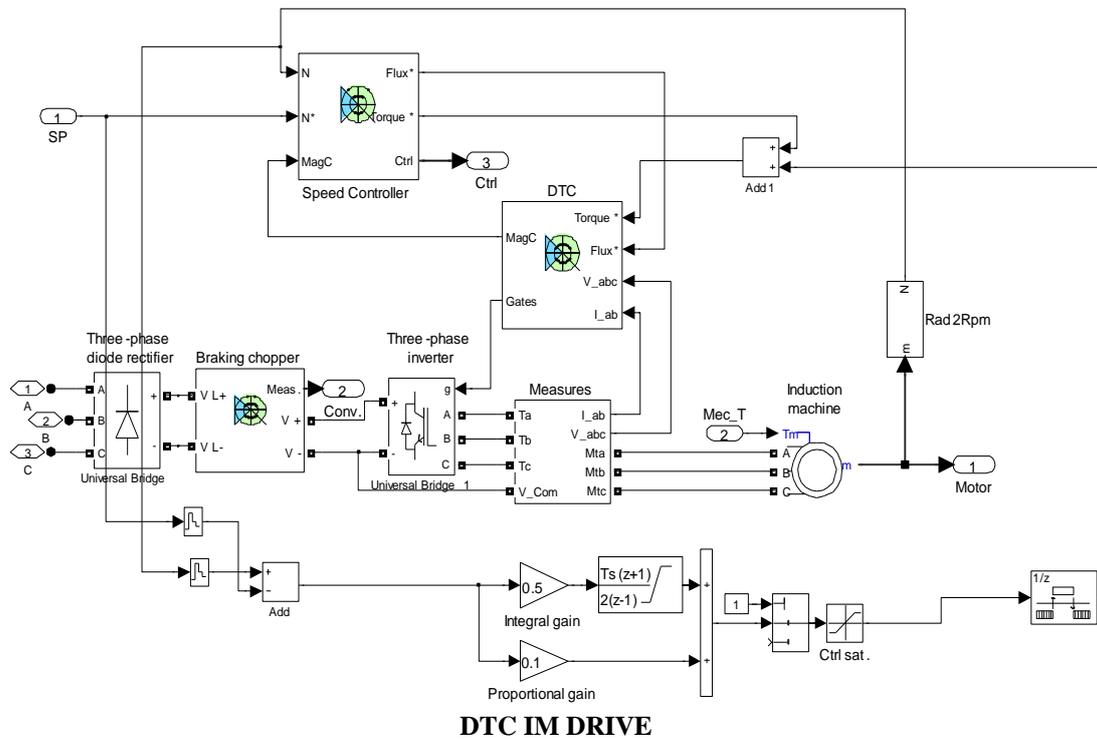
DTC WITH OUT SVM



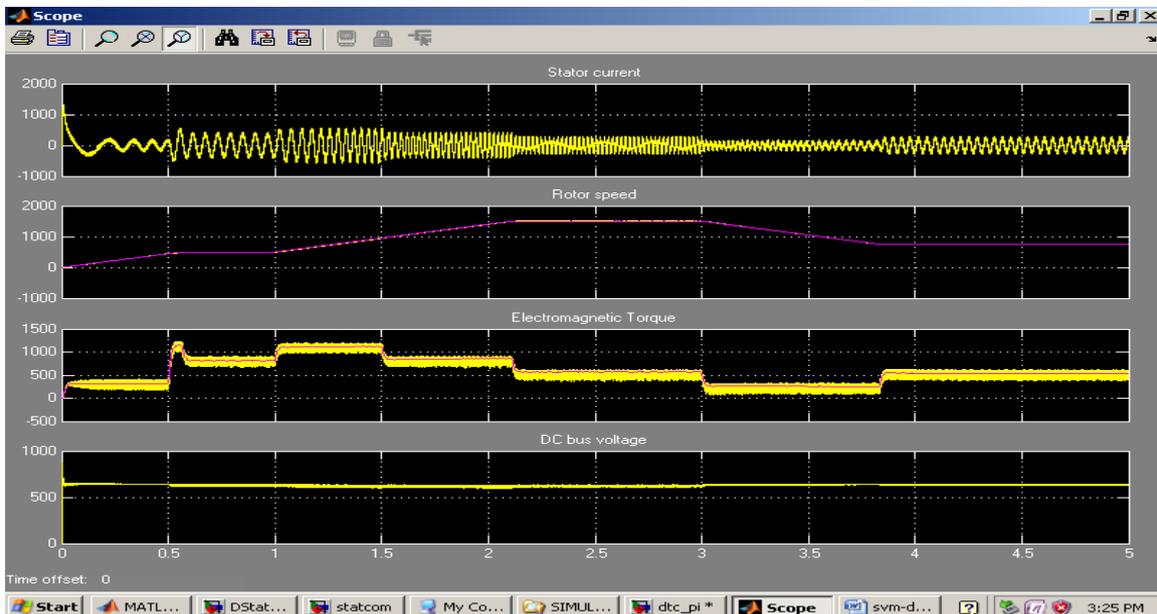
Discrete,
Ts= 1e-006 s.

DTC Induction Motor Drive during speed regulation

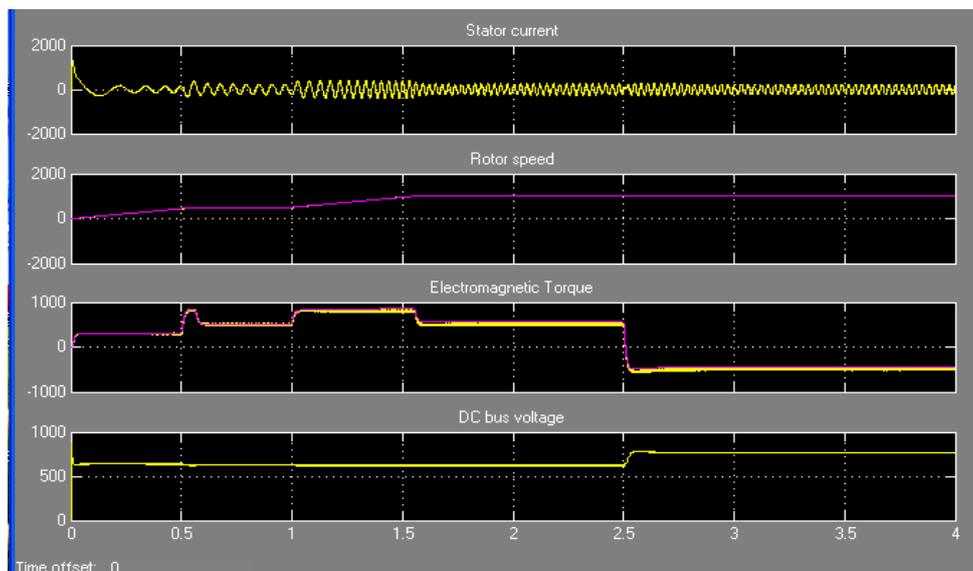
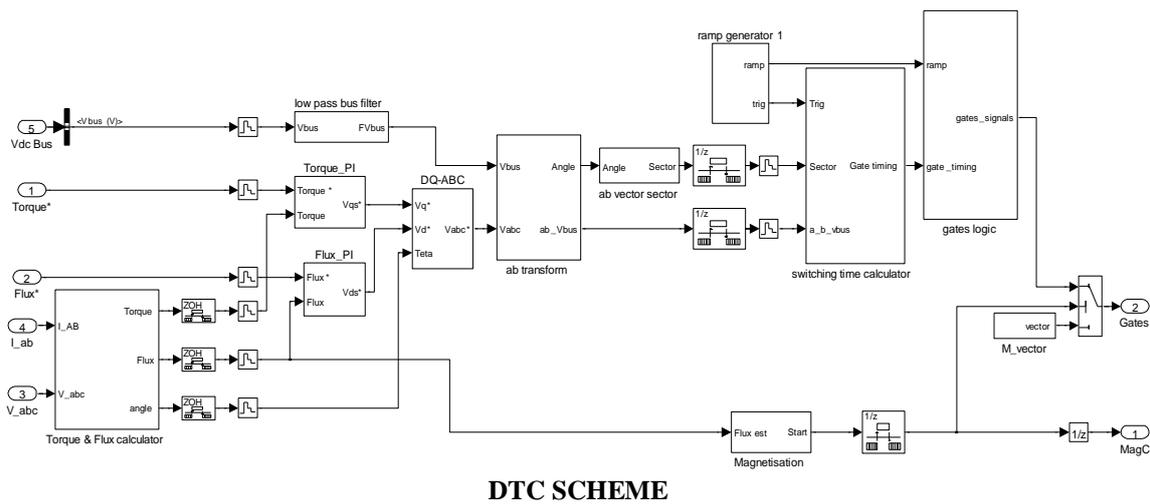
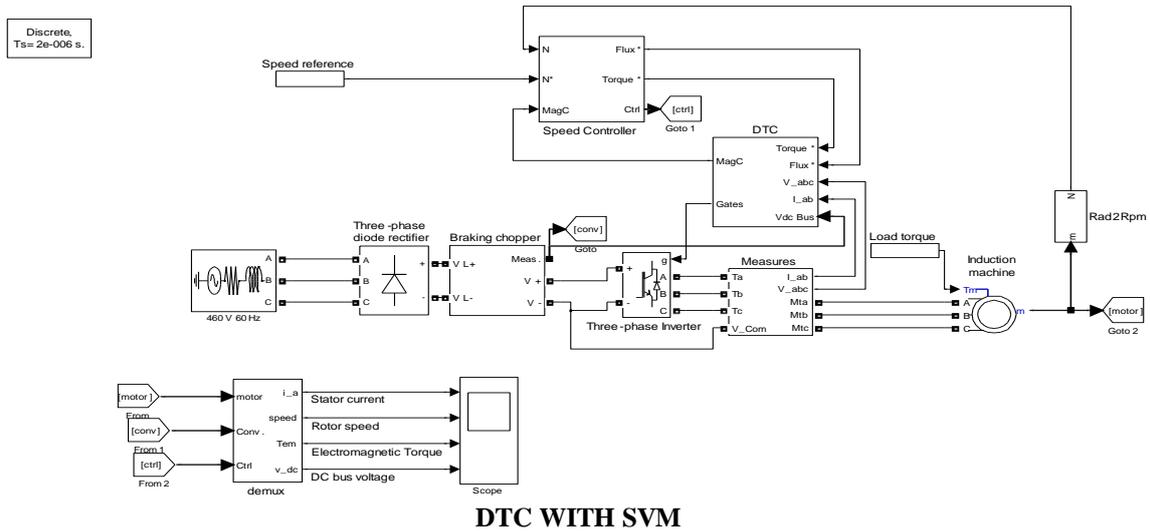
More Info



DTC IM DRIVE



OUTPUT RESPONSE WITH OUT SVM



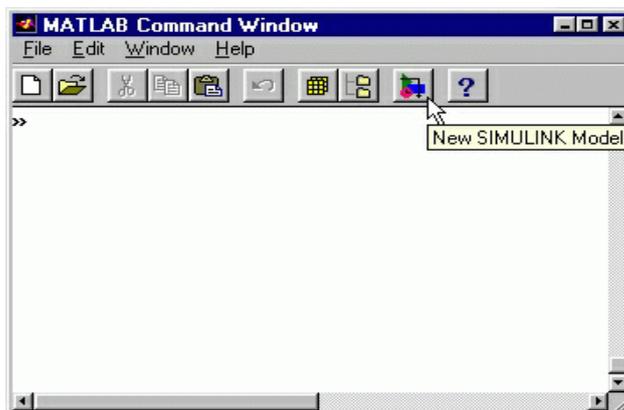
SIMULINK:

Simulink is a platform for multinomial simulation and Model-Based Design for dynamic systems. It provides an interactive graphical environment and a customizable set of block libraries, and can be extended for specialized applications.

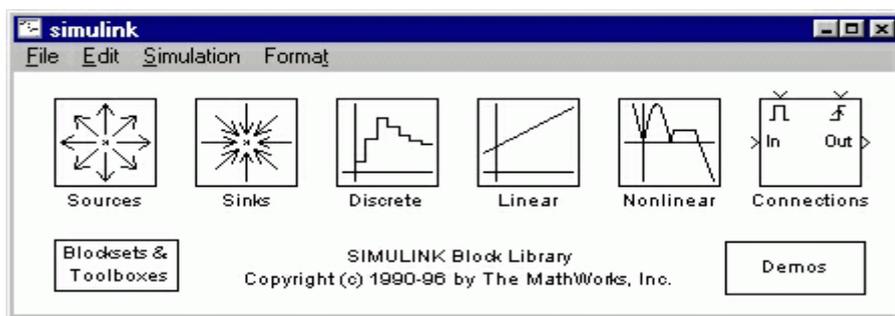
Simulink is a graphical extension to MATLAB for modeling and simulation of systems. In Simulink, systems are drawn on screen as block diagrams. Many elements of block diagrams are available, such as transfer functions, summing junctions, etc., as well as virtual input and output devices such as function generators and oscilloscopes. Simulink is

integrated with MATLAB and data can be easily transferred between the programs. In these tutorials, we will apply Simulink to the examples from the MATLAB tutorials to model the systems, build controllers, and simulate the systems. Simulink is supported on UNIX, Macintosh, and Windows environments; and is included in the student version of MATLAB for personal computers.

Simulink is started from the MATLAB command prompt by entering the following command: `simulink`
 Alternatively, you can hit the New Simulink Model button at the top of the MATLAB command window as shown below:



When it starts, Simulink brings up two windows. The first is the main Simulink window, which appears as:



The second window is a blank, untitled, model window. This is the window into which a new model can be drawn.

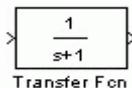
Basic Elements

There are two major classes of items in Simulink: **blocks** and **lines**. Blocks are used to generate, modify, combine, output, and display signals. Lines are used to transfer signals from one block to another.

Blocks: There are several general classes of blocks:

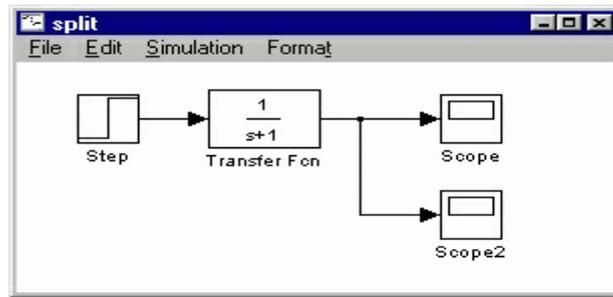
- Sources: Used to generate various signals
- Sinks: Used to output or display signals
- Discrete: Linear, discrete-time system elements (transfer functions, state-space models, etc.)
- Linear: Linear, continuous-time system elements and connections (summing junctions, gains, etc.)
- Nonlinear: Nonlinear operators (arbitrary functions, saturation, delay, etc.)
- Connections: Multiplex, Demultiplex, System Macros, etc.

Blocks have zero to several input terminals and zero to several output terminals. Unused input terminals are indicated by a small open triangle. Unused output terminals are indicated by a small triangular point. The block shown below has an unused input terminal on the left and an unused output terminal on the right.



Lines

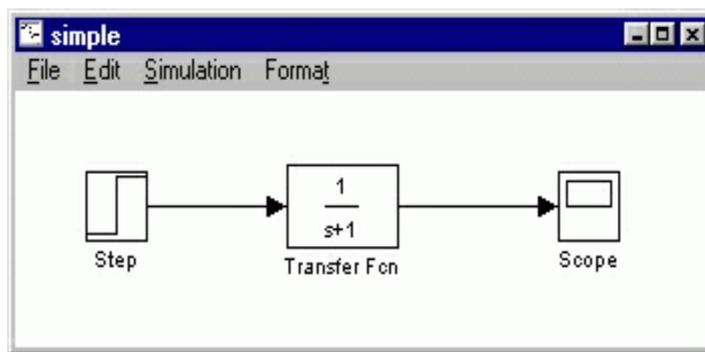
Lines transmit signals in the direction indicated by the arrow. Lines must always transmit signals from the output terminal of one block to the input terminal of another block. On exception to this is a line can tap off of another line, splitting the signal to each of two destination blocks, as shown below (click the figure to download the model file called split.mdl).



Lines can never inject a signal *into* another line; lines must be combined through the use of a block such as a summing junction.

A signal can be either a scalar signal or a vector signal. For Single-Input, Single-Output systems, scalar signals are generally used. For Multi-Input, Multi-Output systems, vector signals are often used, consisting of two or more scalar signals. The lines used to transmit scalar and vector signals are identical. The type of signal carried by a line is determined by the blocks on either end of the line.

Simple Example

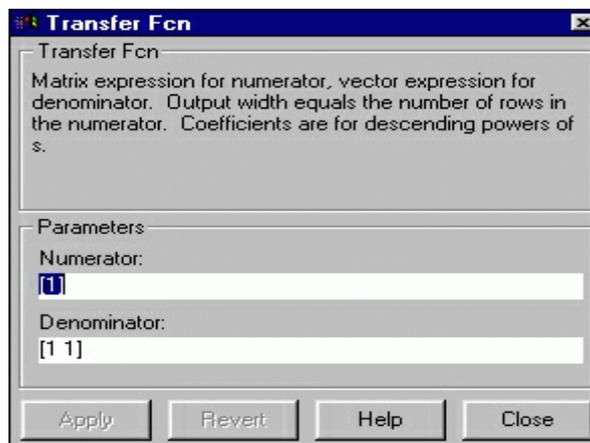


The *simple* model (from the model file section) consists of three blocks: Step, Transfer Fcn, and Scope. The Step is a **source block** from which a step input signal originates. This signal is transferred through the **line** in the direction indicated by the arrow to the Transfer Function **linear block**. The Transfer Function modifies its input signal and outputs a new signal on a line to the Scope. The Scope is a **sink block** used to display a signal much like an oscilloscope.

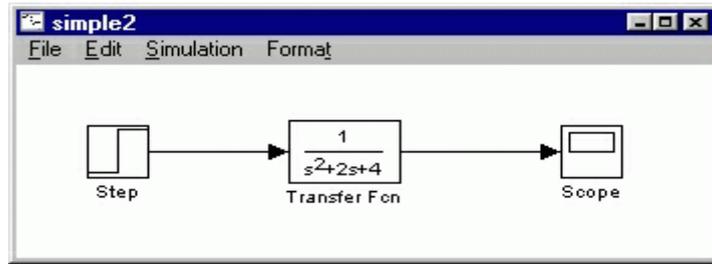
There are many more types of blocks available in Simulink, some of which will be discussed later. Right now, we will examine just the three we have used in the *simple* model.

Modifying Blocks

A block can be modified by double-clicking on it. For example, if you double-click on the "Transfer Fcn" block in the *simple* model, you will see the following dialog box.

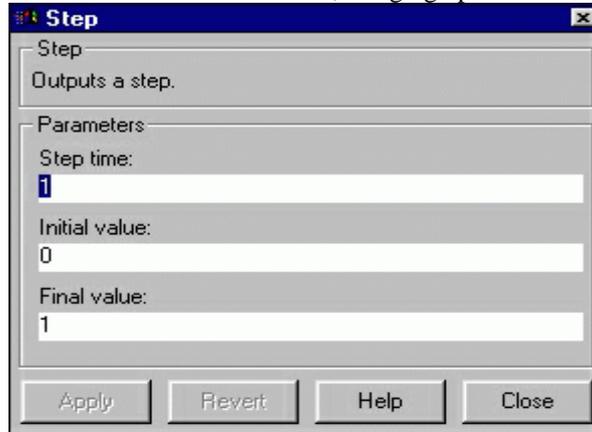


This dialog box contains fields for the numerator and the denominator of the block's transfer function. By entering a vector containing the coefficients of the desired numerator or denominator polynomial, the desired transfer function can be entered. For example, to change the denominator to s^2+2s+1 , enter the following into the denominator field: [1 2 1] and hit the close button, the model window will change to the following,

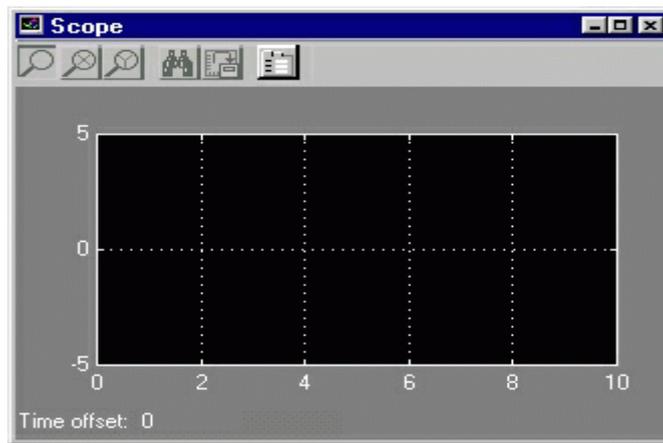


which reflects the change in the denominator of the transfer function.

The "step" block can also be double-clicked, bringing up the following dialog box.



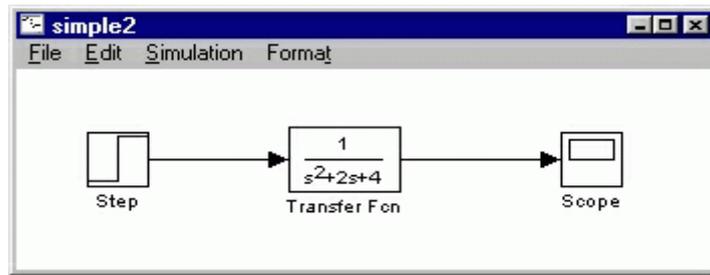
The default parameters in this dialog box generate a step function occurring at time=1 sec, from an initial level of zero to a level of 1. (in other words, a unit step at t=1). Each of these parameters can be changed. Close this dialog before continuing. The most complicated of these three blocks is the "Scope" block. Double clicking on this brings up a blank oscilloscope screen.



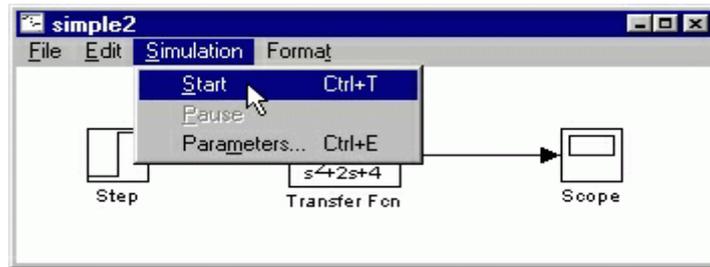
When a simulation is performed, the signal which feeds into the scope will be displayed in this window. Detailed operation of the scope will not be covered in this tutorial. The only function we will use is the auto scale button, which appears as a pair of binoculars in the upper portion of the window.

Running Simulations

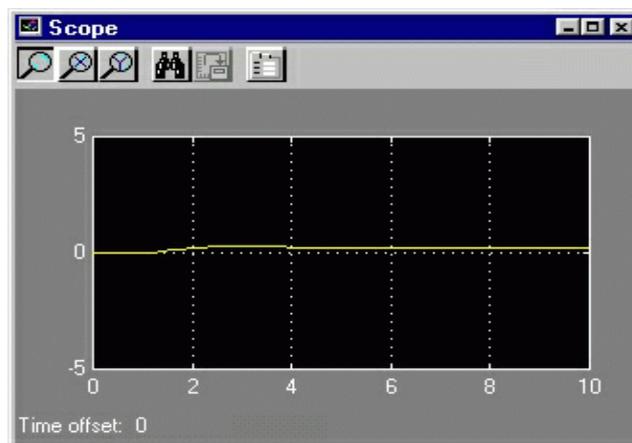
To run a simulation, we will work with the following model file:



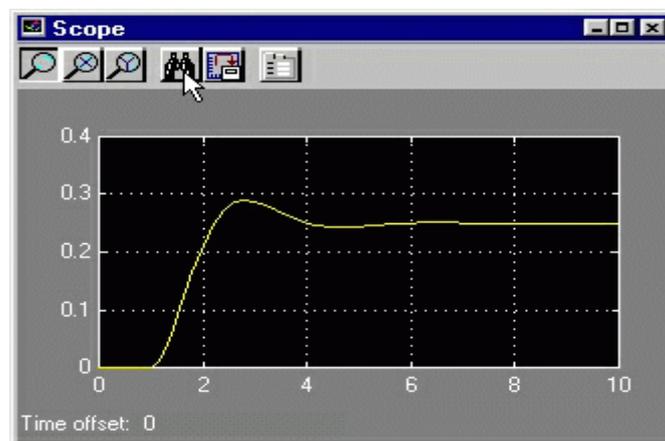
Before running a simulation of this system, first open the scope window by double-clicking on the scope block. Then, to start the simulation, either select **Start** from the **Simulation** menu (as shown below) or hit Ctrl-T in the model window.



The simulation should run very quickly and the scope window will appear as shown below.

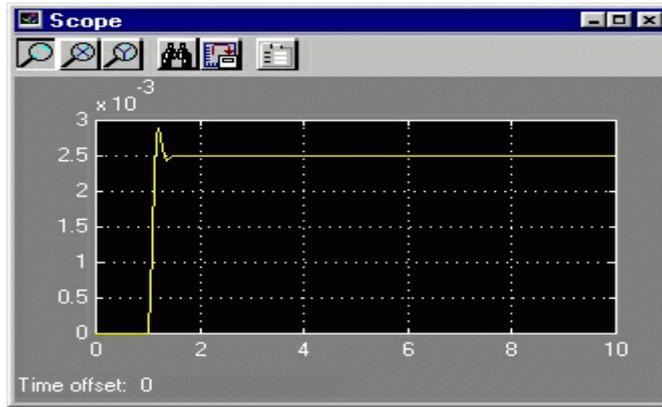


Note that the simulation output (shown in yellow) is at a very low level relative to the axes of the scope. To fix this, hit the auto scale button (binoculars), which will rescale the axes as shown below.

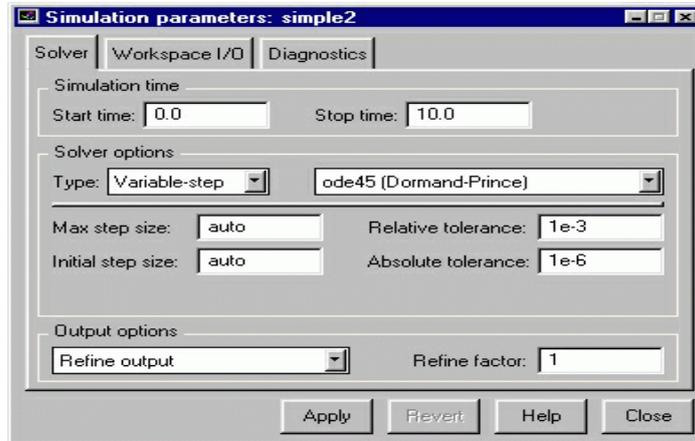


Note that the step response does not begin until $t=1$. This can be changed by double-clicking on the "step" block. Now, we will change the parameters of the system and simulate the system again. Double-click on the "Transfer Fcn" block in the model window and change the denominator to
 $[1 \ 20 \ 400]$

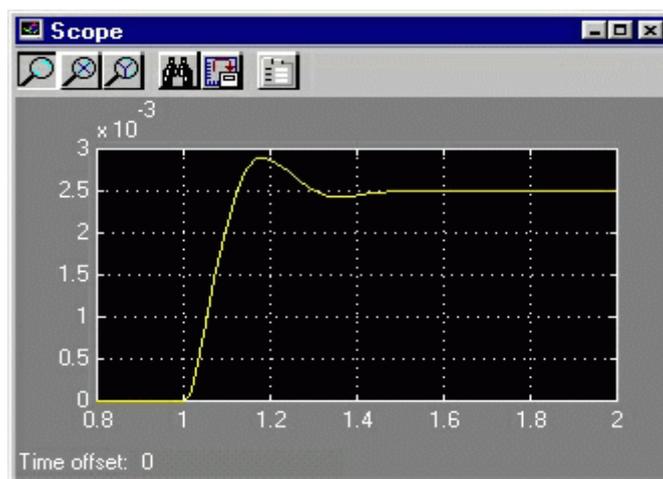
Re-run the simulation (hit Ctrl-T) and you should see what appears as a flat line in the scope window. Hit the auto scale button, and you should see the following in the scope window.



Notice that the auto scale button only changes the vertical axis. Since the new transfer function has a very fast response, it compressed into a very narrow part of the scope window. This is not really a problem with the scope, but with the simulation itself. Simulink simulated the system for a full ten seconds even though the system had reached steady state shortly after one second. To correct this, you need to change the parameters of the simulation itself. In the model window, select **Parameters** from the **Simulation** menu. You will see the following dialog box.



There are many simulation parameter options; we will only be concerned with the start and stop times, which tell Simulink over what time period to perform the simulation. Change **Start time** from 0.0 to 0.8 (since the step doesn't occur until $t=1.0$). Change **Stop time** from 10.0 to 2.0, which should be only shortly after the system settles. Close the dialog box and rerun the simulation. After hitting the auto scale button, the scope window should provide a much better display of the step response as shown below.



IX. CONCLUSION

In classical DTC, as the torque ripple is maintained within hysteresis band, switching frequency changes with speed. Moreover, the torque ripple is important problem at low speed. So using constant switching frequency a desired torque ripple can be achieved at low speeds where it really matters. The torque ripple for this SVM-DTC is significantly improved and switching frequency is maintained constant. Numerical simulations have been carried out showing the advantages of the SVM-DTC method with respect to the conventional DTC.

REFERENCES

- [1] S. Jianguo and C. Quanshi, "Research of electric vehicle IM controller based on space vector modulation direct torque control," *Proceedings of the Eighth International Conference on Electrical Machines and Systems, ICEMS 2005*, vol. 2, pp. 1617-1620, 2005.
- [2] D. Ocean, L. Romeral, J.A. Ortega, J. Cusido, and A. Garcia, "Discrete space vector modulation applied on a PMSM motor," *12th International Power Electronics and Motion Control Conference, EPE-PEMC 2006* p. 320-325, 2006.
- [3] H.R. Keyhani, M.R. Zolghadri, and A. Homaifar, "An extended and improved discrete space vector modulation direct torque control for induction motors," *35th Annual IEEE Power Electronics Specialists Conference*, Germany, pp. 3414-3420, 2004.
- [4] R.H. Ahmad, G.G. Karady, T.D. Blake, and P. Pinewski, "Comparison of space vector modulation techniques based on performance indexes and hardware implementation," *23rd International Conference on Industrial Electronics, Control and Instrumentation, IECON 97*, vol. 2, p. 682-687, 1997.
- [5] I. Takashi and T. Noguchi, "A new quick-response and high-efficiency control of an induction motor," *IEEE Trans. Industry Applications*, vol. A-22, no.5, pp. 820-827, 1986.
- [6] M. Depenbrock, "Direct self control (DSC) of inverter-fed induction machines," *IEEE Trans. Power Electronics*, vol. 3, no. 4, pp. 420-429, 1988.
- [7] M. Jin, J. Qiu, C. Shi, and R. Lin, "A fuzzy DTC method with a SVM defuzzification to permanent magnet synchronous machine," *The 30th Annual Conference of the IEEE Industrial Electronics Society, Korea* p. 3196-3199, 2004
- [8] M. Vasudevan and R. Arumugam, "New direct torque control scheme of induction motor for electric vehicles," *5th Asian Control Conference*, pp. 377-383, 2004.
- [9] A. Kumar, B.G. Fernandes, and K. Chatterjee, "Simplified SVPWM-DTC of 3-phase induction motor using the concept of imaginary switching times," *The 30th Annual Conference of the IEEE Industrial Electronics Society, Korea*, pp. 341-346, 2004.
- [10] U. Senthil and B.G. Fernandes, "Hybrid space vector pulse width modulation based direct torque controlled induction motor drive," *IEEE 4th Annual Power Electronics Specialist Conference, PESC'02*, vol. 3, p. 1112-1117, 2003.
- [11] . Chen, K.L. Fang, and Z.F. Hu, "A scheme of fuzzy direct torque control for induction machine," *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou*, pp. 03-807, 2005.
- [12] H.I. Okumus, "Improved direct torque control of induction machine drives," *PhD Thesis, University of Bristol, UK*, July 2001.
- [13] L. Tang, M.F. Rahman, "A new direct torque control strategy for flux and torque ripple reduction for induction motors drive by using space vector modulation," *IEEE 32nd Annual Power Electronics Specialists Conference, PESC 2001*, vol. 3, pp. 1440-1445, 2001.
- [14] D. Swierczynski, M.P. Kazmierkowski, and F. Blaabjerg, "DSP based direct torque control of permanent magnet synchronous motor (PMSM) using space vector modulation (DTC-SVM)," *Proceedings of the 2002 IEEE International Symposium on Industrial Electronics, ISIE 2002*, vol. , pp. 723-727, 2002.

Iterative Multivariate Interpolation for Low Complexity Reed-Solomon Codes

N. Sireesha¹, V. Prasanth²

¹ PG Scholar, ECE, Pragati Engineering College, JNTUK, AP, INDIA

² Associate Professor, ECE, Pragati Engineering College, JNTUK, AP, INDIA

ABSTRACT : The algebraic soft-decoding (ASD) of Reed-Solomon (RS) codes provides significant coding gain over hard-decision decoding with polynomial complexity. In order to reduce the complexity in this paper, high-throughput interpolator architecture for soft-decision decoding of Reed-Solomon (RS) codes based on low-complexity chase (LCC) decoding is presented. An efficiency is low, in terms of area-delay product, has been achieved by an LCC decoder, by using the proposed interpolator architecture, over the best of the previously reported architectures for an RS(255,239) code with eight test vectors. We have implemented the proposed interpolator in CYCLONE III FPGA which provides 1.3 Gb/s throughput.

Keywords: Algebraic soft-decision decoding, interpolation, low-complexity chase (LCC), low latency, Nielson's algorithm, Reed-Solomon (R-S) codes.

I. INTRODUCTION

The Reed-Solomon codes (RS codes) are non binary cyclic codes with code symbols from a Galois field. We want to transmit a message f . The bits of the message can be grouped in $\log_2(q)$ -bit symbols chosen from the finite field with q elements, $GF(q)$. An (n, k) Reed-Solomon code over $GF(q)$ represents the k -symbol message, $f = (f_0, f_1, f_2, \dots, f_{k-1})$ by an n -symbol codeword, $c = (c_0, c_1, c_2, \dots, c_{k-1}, \dots, c_{n-1})$, where $n > k$ and usually $n = q - 1$. The k symbols of the message f can be considered to be the coefficients of the up to degree $(k - 1)$ univariate message polynomial.

$$f(x) = f_0 + f_1x + f_2x^2 + \dots + f_{k-1}x^{k-1}. \quad (1)$$

We use the classical view of Reed-Solomon codes taken from the original definition, with this evaluation map encoding method, a codeword is formed by evaluating the message polynomial $f(x)$ at n elements of $GF(q)$. If the set of evaluation elements is $X = \{x_0, x_1, \dots, x_{n-1}\}$, the codeword c .

$$c = (f(x_0), f(x_1), \dots, f(x_{n-1})), \quad x_i \in X. \quad (2)$$

We will always assume that $n = q - 1$ and the set of evaluation elements X is the set of nonzero elements of $GF(q)$:

$$X = \{x_0, x_1, x_2, \dots, x_{n-1}\} = \{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\} \quad (3)$$

where α is a primitive n 'th root of unity. The evaluation map encoding method is useful because, it provides insight leading to interpolation-based decoding algorithms.

Guruswami-Sudan algorithm

An interpolation-based decoder takes the point of view that a codeword is a message polynomial evaluated at points in

a finite field and uses polynomial interpolation to try to reconstruct that polynomial. The Guruswami-Sudan (GS)

algorithm [5] is an interpolation-based list decoder for Reed-Solomon codes. To describe the algorithm, we will first need to review some notation and facts about bivariate polynomials, which are the basic data structures in the algorithm. Consider the bivariate polynomial [14] with coefficients chosen from a finite field:

$$P(x, y) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} P_{a,b} x^a y^b \in GF(q)[x, y]$$

Consider the received word $y = c + e$, where e is an error vector with components drawn from $GF(q)$. Since each component of c was generated by evaluating $f(x)$ at a unique value of $x \in X$, a unique x_i can be associated with each received $y_i \in GF(q)$ to form the list of points, $L = \{(x_0, y_0), (x_1, y_1), \dots, (x_{n-1}, y_{n-1})\}$. If there is no noise ($e = 0$), then $y_i = f(x_i)$, $0 \leq i < n$, and a bivariate polynomial, $P(x, y) = y - f(x)$, passes through all the points in L with a multiplicity of one. This suggests that an interpolation-based approach can be used to decode Reed-Solomon codes. In the presence of noise ($e \neq 0$), the interpolation polynomial will pass through some points that are not part of the codeword. The GS algorithm ensures that under certain conditions, the codeword polynomial "lives inside" the interpolation polynomial [2, 3].

II. INTERPOLATION

The GS algorithm is an interpolation-based list decoder with two main steps [11].

1. Interpolation Step: Given the set of points L and a positive integer m , compute $P(x, y)$ of $GF(q)[x, y] \setminus \{0\}$ of minimal $(1, k - 1)$ -weighted degree that passes through all the points in L with multiplicity at least m .

2. Factorization Step: Given the interpolation polynomial $P(x, y)$, identify all the factors of $P(x, y)$ of the form $y - f(x)$ with $\deg f(x) < k$. The output of the algorithm is a list of the codewords that correspond to these factors.

A complete factorization of $P(x, y)$ is not necessary since we are just looking for linear y -roots of degree $< k$. An appropriate root-finding algorithm is given. The multiplicity, m , functions as a user-selectable complexity parameter. The error-correcting ability of the GS algorithm increases as the value of m increases.

Primitive polynomials are of interest here because they are used to define the Galois field. A popular choice for a primitive polynomial is:

$$p(x) = x^8 + x^7 + x^2 + x^1 + 1 \quad (4)$$

This is also known as the 0x87 polynomial, corresponding to the binary representation of the polynomial's coefficients excluding the MSB (i.e. 10000111). This specific polynomial is used in the CCSDS specification for a RS (255, 223). In $GF(2^8)$ there are 16 possible primitive polynomials.

A. FINITE FIELDS

In order to understand the encoding and decoding principles of nonbinary codes, such as Reed-Solomon (R-S) codes, it is necessary to venture into the area of finite fields known as Galois Fields (GF). For any prime number, p, there exists a finite field denoted GF(p) that contains p elements. It is possible to extend GF(p) to a field of pm elements, called an extension field of GF(p), and denoted by GF(pm), where m is a nonzero positive integer. Note that GF(pm) contains as a subset the elements of GF(p). Symbols from the extension field GF(2m) are used in the construction of Reed-Solomon (R-S) codes.

The binary field GF(2) is a subfield of the extension field GF(2m), in much the same way as the real number field is a subfield of the complex number field. Besides the numbers 0 and 1, there are additional unique elements in the extension field that will be represented with a new symbol α . Each nonzero element in GF(2m) can be represented by a power of α . An infinite set of elements, F, is formed by starting with the elements {0, 1, α }, and generating additional elements by progressively multiplying the last entry by α .

$$F = \{0, 1, \alpha, \alpha^2, \dots, \alpha^j, \dots\} = \{0, \alpha^0, \alpha^1, \alpha^2, \dots, \alpha^j, \dots\} \quad (5)$$

To obtain the finite set of elements of GF(2m) from F, a condition must be imposed on F so that it may contain only 2m elements and is closed under multiplication. The condition that closes the set of field elements under multiplication is characterized by the irreducible polynomial shown below:

$$\alpha^{2m-1} + 1 = 0 \quad (6)$$

or equivalently Using this polynomial constraint, any field element that has a power equal to or greater than 2m - 1 can be reduced to an element with a power less than 2m - 1, as follows:

$$\alpha^{2m+n} = \alpha^{2m-1} \alpha^{n+1} = \alpha^{n+1} \quad (7)$$

B. REVIEW OF LCC DECODING OF RS CODES

The Reed-Solomon (R-S) codes are particularly useful for burst-error correction; that is, they are effective for channels that have memory. Also, they can be used efficiently on channels where the set of input symbols is large. An interesting feature of the R-S code is that as many as two information symbols can be added to an R-S code of length n without reducing its minimum distance.

For R-S codes, error probability is an exponentially decreasing function of block length, n, and decoding complexity is proportional to a small power of the block length. The R-S codes are sometimes used in a concatenated arrangement. In such a system, an inner convolution decoder first provides some error control by operating on soft-decision demodulator outputs; the convolutional decoder then presents hard-decision data to the outer Reed-Solomon decoder, which further reduces the probability of error.

III. SYNDROME COMPUTATION

Reed-Solomon codes are non-binary cyclic codes with symbols made up of m-bit Sequences, where m is any positive integer having a value greater than 2. R-S (n, k) codes on m-bit symbols exist for all n and k for which

$$0 < k < n < 2m + 2 \quad (8)$$

Where k is the number of data symbols being encoded, and n is the total number of code symbols in the encoded block. For the most conventional R-S (n, k) code,

$$(n, k) = (2m - 1, 2m - 1 - 2t) \quad (8a)$$

Where t is the symbol-error correcting capability of the code, and n - k = 2t is the number of parity symbols. An extended R-S code can be made up with n = 2m or n = 2m + 1, but not any further. In 2t syndromes are calculated by evaluating received polynomial for powers of q.

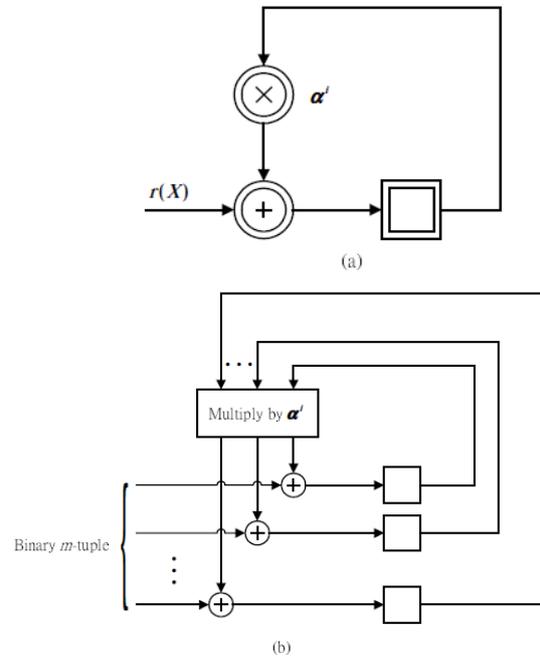


Figure 1: Syndrome computation circuits for Reed-Solomon codes: (a) over GF(2m); (b) in binary form.

Assume the transmitted code vector is

$$t(X) = t_0 + t_1X + t_2X^2 + \dots + t_{n-1}X^{n-1},$$

and the received vector is

$$r(X) = r_0 + r_1X + r_2X^2 + \dots + r_{n-1}X^{n-1}.$$

The first step in decoding a Reed-Solomon code is to calculate the 2t syndrome components as:

$$S_0 = r(a^0) = r_0 + r_1 + r_2 + \dots + r_{n-1}$$

$$S_1 = r(a^1) = r_0 + r_1(a) + r_2(a)^2 + \dots + r_{n-1}(a)^{n-1}$$

$$S_2 = r(a^2) = r_0 + r_1(a^2) + r_2(a^2)^2 + \dots + r_{n-1}(a^2)^{n-1}$$

$$S_{2t-1} = r(a^{2t-1}) = r_0 + r_1(a^{2t-1}) + r_2(a^{2t-1})^2 + \dots + r_{n-1}(a^{2t-1})^{n-1}.$$

The syndrome polynomial is

$$S(X) = S_0 + S_1X + S_2X^2 + \dots + S_{2t-1}X^{2t-1}.$$

The second step in decoding a Reed-Solomon code is to find the error location polynomial L(X) and the error evaluation polynomial W(X).

The error location polynomial is

$$L(X) = 1 + L_1X + L_2X^2 + \dots + L_eX^e,$$

and the error evaluation polynomial is

$$W(X) = W_0 + W_1X + W_2X^2 + \dots + W_{e-1}X^{e-1},$$

where e is the number of errors. The error location polynomial and the error evaluation polynomial are related to the syndrome polynomial through the key equation is

$$L(X)S(X) = W(X) \text{ mod } X^{2t}.$$

The popular iterative Berlekamp - Macey algorithm is

used to solve for $L(X)$ and $M(X)$. The last step in decoding a Reed-Solomon code is to find the error location and the error value. The error location is obtained using Chan's searching algorithm. Basically X is substituted with a^n in $L(X)$ for all possible n in a code to find the root of $L(X)$. The inverse of the root of the error location polynomial is the error position. After an error location is found, the error value is calculated via Forney's error evaluation algorithm. Once the error value is found, it is added to the corrupted symbol to correct the error.

IV. IMPLEMENTATIONS

The implementations below can be customized to work with other RS (n, k) codes to yield similar results in performance.

Optimized Software Implementation: The pure software implementation is dominated computationally by multiplication over a finite field (Galois Field multiplication). The encoder requires 71,181 cycles per codeword on a MIPS32 processor and the decoder requires 66,045 cycles.

Scalar GF Multiply Support: This is the simplest form of VOCAL's hardware acceleration. The Scalar GF Multiply Support extends the capabilities of the MIPS32 processor by taking advantage of MIPS Technologies CorExtend capability to decrease the number of cycles to 23,305 cycles to encode and 9,174 cycles per codeword to decode on the MIPS32 processor.

SIMD GF Multiply Support: The SIMD GF Multiply Support requires 128 bytes of local ROM Memory, but increases the performance to 3,918 cycles per megabit to encode and 3,078 cycles per codeword to decode. RS Encode Kernel. The RS Encode Kernel uses 1024 bytes of local ROM memory to encode. The number of cycles to process a codeword on a MIPS32 CPU falls to 2,702 cycles for encoding and decoding only consumes 828 cycles with this implementation.

V. PROPOSED ARCHITECTURE

Methodologies are the principles and explanations of High-Throughput Interpolator Architecture for Low-Complexity Chase Decoding of RS Codes. And here we have Five types of modules are used

A. REGISTERS

The Shift Register is used for data storage or data movement and are used in calculators or computers to store data such as two binary numbers before they are added together, or to convert the data from either a serial to parallel or parallel to serial format. The individual data latches that make up a single shift register are all driven by a common clock signal making them synchronous devices.

B. MULTIPLEXERS

A $2n$ -to-1 multiplexer sends one of $2n$ input lines to a single output line. A multiplexer has two sets of inputs: $2n$ data input lines, n select lines, to pick one of the $2n$ data inputs. The mux output is a single bit, which is one of the $2n$ data inputs. A $2n$ -to-1 multiplexer routes one of $2n$ input lines to a single output line. Just like decoders, muxes are common enough to be supplied as stand-alone devices for use in modular designs. Muxes can implement arbitrary functions.

Smaller muxes can be combined to produce larger ones. It can add active-low or active-high enable inputs. As always, we use truth tables and Boolean algebra to analyze things. Tune in tomorrow as we start to discuss how to build circuits to do arithmetic.

C. D-FLIP-FLOP

There are some circuits that are not quite as straight forward as the gate circuits. However, we still need to learn about circuits that can store and remember information. They're the kind of circuits that are used in computers to store program information - RAM memory. The combination of two flip-flops constitutes a D-type flip-flop. That's D because the output of the flip-flop is delayed by the time of one clock pulse. Set a value for the data and pulse the clock ON and OFF. We'll find a copy of the data appearing at the output on the trailing edge of the clock pulse. Now, if we consider the combination of two flip-flops as a unit, we have a D flip-flop. It's called a D flip-flop because it delays the signal. The signal appears at the output of the circuit delayed by the time of one clock pulse.

D. GF (2^8) MULTIPLIER

Galois Field Theory (GFT) deals with numbers that are binary in nature, have the properties of a mathematical "field," and are finite in scope. Although some Galois computations don't exist in ordinary mathematics, many Galois operations match those of regular math. Addition (Ex-Or) and multiplication are common Galois operations, and logarithms, particularly, are handy for checking multiplication results. For over 40 years, Galois Field multipliers have been used both for coding theory and for cryptography. Both areas are complex, with similar needs, and both deal with fixed symbolic alphabets that neatly fit the extended Galois Field model.

The Finite Field GF (2^8):

The case in which n is greater than one is much more difficult to describe. In cryptography, one almost always takes p to be 2 in this case. This section just treats the special case of $p = 2$ and $n = 8$, that is, $GF(2^8)$, because this is the field used by the new U.S. Advanced Encryption Standard (AES). The AES works primarily with bytes (8 bits), represented from the right as: $b7b6b5b4b3b2b1b0$. The 8-bit elements of the field are regarded as polynomials with coefficients in the field Z_2 : $b7x^7 + b6x^6 + b5x^5 + b4x^4 + b3x^3 + b2x^2 + b1x + b0$. The field elements will be denoted by their sequence of bits, using two hex digits.

Multiplication in GF (2^8):

Multiplication in this field is much more difficult and harder to understand, but it can be implemented very efficiently in hardware and software. The first step in multiplying two field elements is to multiply their corresponding polynomials just as in beginning algebra (except that the coefficients are only 0 or 1, and $1 + 1 = 0$ makes the calculation easier, since many terms just drop out). The result would be up to a degree 14 polynomial -- too big to fit into one byte. A finite field now makes use of a fixed degree eight irreducible polynomial (a polynomial that cannot be factored into the product of two simpler polynomials). For the AES the polynomial used is the following (other

polynomials could have been used): $m(x) = x^8 + x^4 + x^3 + x + 1 = 0x11b$ (hex). The intermediate product of the two polynomials must be divided by $m(x)$. The remainder from this division is the desired product. This sounds hard, but is easier to do by hand than it might seem (though error-prone). To make it easier to write the polynomials down, adopt the convention that instead of $x^8 + x^4 + x^3 + x + 1$ just write the exponents of each non-zero term. (Remember that terms are either zero or have a 1 as coefficient.)

E. GF (2⁸) ADDER

To add two field elements, just add the corresponding polynomial coefficients using addition in Z_2 . Here addition is modulo 2, so that $1 + 1 = 0$, and addition, subtraction and exclusive-or are all the same. The identity element is just zero: 00000000 (in bits) or 0x00 (hex).

SIMULATION RESULTS

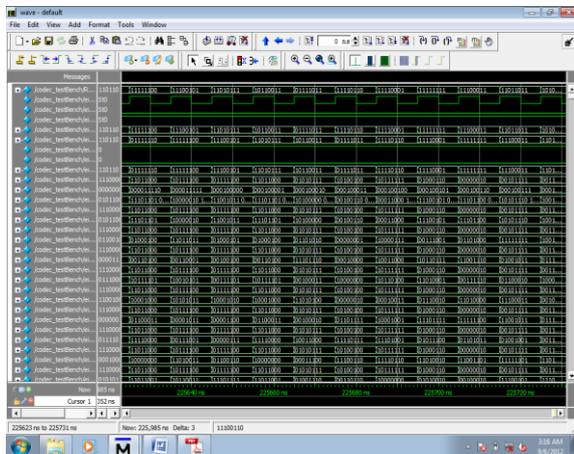


Figure 2: Simulated Output.

AREA UTILIZATION REPORT

Flow Summary	
Flow Status	Successful - Thu Sep 06 03:20:19 2012
Quartus II Version	11.0 Build 208 07/03/2011 SP 1 SJ Web Edition
Revision Name	test
Top-level Entity Name	RS8FreqDecode
Family	Cyclone III
Total logic elements	7,012 / 15,408 (46 %)
Total combinational functions	7,012 / 15,408 (46 %)
Dedicated logic registers	2,479 / 15,408 (16 %)
Total registers	2479
Dedicated logic registers	21 / 169 (12 %)
Total virtual pins	0
Total memory bits	2,048 / 516,096 (< 1 %)
Embedded Multiplier 9-bit elements	0 / 112 (0 %)
Total PLLs	0 / 4 (0 %)
Device	EP3C16F256C6
Timing Models	Final

Figure 3: Flow Summary Report

V. CONCLUSIONS

The proposed one is a hardware implementation of modified Nielson's algorithm, which works with a different scheduling, leads limited growth of the polynomials and shares the common interpolation points, for reducing the latency of interpolation. Based on the proposed modified Nielson's algorithm, we have derived a multivariate interpolator architecture. This will reduce the number of

iterations required for LCC decoder. Using our low-latency interpolator is found to be at least 45% more efficient in terms of area-delay product over the best of previous works. This architecture has been implemented in a CYCLONE-III FPGA device which provides a throughput of 1.3 Gb/s .

REFERENCES

- [1] F. García-Herrero, M. J. Canet, J. Valls, and P. K. Meher "High-Throughput Interpolator Architecture for Low-Complexity Chase Decoding of RS Codes" *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 3, pp. 568–573, Mar. 2011
- [2] R. Koetter and A. Vardy, "Algebraic soft-decision decoding of Reed–Solomon codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2809–2825, Nov. 2003.
- [3] A. Ahmed, N. R. Shanbhag, and R. Koetter, "An architectural comparison of Reed–Solomon soft-decoding algorithm," *Signals, Syst. Comput.*, pp. 912–916, 2006.
- [4] W. J. Gross, F. R. Kschischang, R. Koetter, and P.G.Gulak, "Architecture and implementation of an interpolation processor for soft-decision Reed–Solomon decoding," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 15, no. 3, pp. 309–318, Mar. 2007.
- [5] X. Zhang, "Reduced complexity interpolation architecture for soft-decision Reed–Solomon decoding," *IEEE Trans. Very Large Scale Integr.(VLSI) Syst.*, vol. 14, no. 10, pp. 1156–1161, Oct. 2006.
- [6] Z. Wang and J. Ma, "High-speed interpolation architecture for softdecision decoding of Reed–Solomon codes," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 9, pp. 937–950, Sep. 2006.
- [7] J. Zhu and X. Zhang, "Efficient VLSI architecture for soft-decision decoding of Reed–Solomon codes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 10, pp. 3050–3062, Nov. 2008.
- [8] J. Bellorado and A. Kavcic, "A low-complexity method for Chase-type decoding of Reed–Solomon codes," *Proc. ISIT*, pp. 2037–2041, Jul. 2006.
- [9] X. Zhang and J. Zhu, "High-throughput interpolation architecture for algebraic soft-decision Reed–Solomon decoding," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 3, pp. 581–591, Mar. 2010.
- [10] J. Zhu, X. Zhang, and Z. Wang, "Backward interpolation architecture for algebraic soft-decision Reed–Solomon decoding," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 17, no. 11, pp. 1602–1615, 2009.
- [11] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. Hoboken, NJ: Wiley, 2004.
- [12] W. J. Gross, F. R. Kschischang, R. Koetter, and P. G. Gulak, "A VLSI architecture for interpolation-based in soft-decision list decoding of Reed–Solomon decoders," *J. VLSI Signal Process.*, vol. 39, no. 1–2, pp. 93–111, 2005.
- [13] F. Parvaresh and A. Vardy, "Multiplicity assignments for algebraic soft-decoding of Reed–Solomon codes," in *Proc. ISIT*, 2003, pp. 205–205.
- [14] X. Zhang, "High-speed VLSI architecture for low-complexity Chase soft-decision Reed–Solomon decoding," in *Proc. Inf. Theory Applic. Workshop*, San Diego, CA, Feb. 2009.
- [15] J. Ma, A. Vardy, and Z. Wang, "Reencoder design for soft-decision decoding of an (255,239) Reed–Solomon code," in *Proc. IEEE Int. Symp. Circuits Syst.*, Island of Kos, Greece, May 2006, pp. 3550–3553.

Numerical Analysis for Effect of Envelop Color of Oil Tank Storage with Floating Roof

Andisheh Tavakoli¹, Mohamadreza Baktash²

¹Department of Material Engineering, Amirkabir University of Technology, Tehran, Iran
Petropars Ltd. Company, Client of Oil and Gas Project

²Department of Mechanic Engineering, K.N.Toosi University of Technology, Tehran, Iran
Namvaran Company, Consultant of Oil and Gas Project

Abstract: The aim of the present work is to determine the evaporation rates from external floating storage tanks and to study the effects of their exterior surface paint on the losses due to the solar irradiation. In this study a numerical scheme has been developed for estimating the time variations of the storage tank temperature and evaporative losses. Considering this fact that the evaporation happens in the fluid surface and the surface temperature is important parameter in this process, investigate the solution for the reducing surface temperature lead to reducing evaporation. One of the methods for reducing surface temperature is reducing effect of solar radiation on the storage tank and for this aim in this study investigated effect of paint color on the evaporation loss. The results show that the absorptivity of the exterior surface paint has considerable effects on tank temperature variations and the evaporative losses accordingly. The value of annual evaporation loss for light color has 170 barrel and for dark color has 370 barrel, in order to difference of evaporation loss between good and bad paint color has 200 barrel. Note that this difference has for one crude oil storage tank and there are about 40 storage tank in Khark Island. Furthermore, the numerical value of monthly averaged evaporation losses have been compared with the estimations based on the API AP-42 standard and good agreement has been observed.

Key words: Storage tank; External floating roof; Evaporative losses; surface Paint; Solar radiation.

I. INTRODUCTION

Design of storage tanks depends on various parameters such as the vapor pressure, storage temperature and pressure, and the toxicity of liquid [1]. The fixing-roof tanks are mainly used for petroleum materials with a vapor pressure less than 1.5 psia [2], while floating-roof tanks are used for petroleum materials with a vapor pressure of 1.12–11.5 psia [1]. An external floating roof tank typically consists of an open topped cylindrical steel shell equipped with a roof that floats on the surface of the storage liquid, which rises and falls as the liquid level changes. Floating roof tanks are equipped with a sealing system, which is attached to the roof perimeter and covers the gap between the roof and the tank wall [3]. The basic designs available for external floating roof rim seals are mechanical shoe seals, liquid-mounted seals, and vapor-mounted that called primary seals [4]. A secondary seal is often used for covering the entire primary seal. The floating roof structure and the sealing system are designed to reduce evaporative losses of the petroleum materials. Evaporative losses from the external floating roof tanks are limited to the losses from the sealing system and roof fittings and any remaining liquids on the tank walls, while the floating roof falls down.

There have been very limited studies related to the storage tank evaporative losses. Wongwises et al. [5] evaluated the gasoline evaporation losses from Thailand storage sites and service stations during refueling and loading. They estimated the total gasoline evaporative losses of about 21,000 tons/year throughout the Thailand. Ramachandran [3], also, investigated the underlying causes of storage tank emissions and analyzed the options of reducing them. Asharif and Zorgani [6] calculated evaporative losses from existing large crude oil storage tanks located in a Libyan oil field and investigated the operating variables including the number of separation stages, operating temperature and pressure of each separator. They concluded that the operation variables of the existing process facilities can be adjusted in order to minimize the losses from storage tanks. Digrado and Thorp [7] compared the evaporation losses between the internal and external floating roofs. They also determined the losses associated with different sealing arrangements based on the American Petroleum Institute (API) standards [8, 9]. Zareie et al. [10] experimentally determined the amount of the volatile organic compounds emitted from an industrial external floating roof tank by monitoring the level of the liquid in the tank and its temperature for a period of 35 days. They also compared their findings with the values computed based on the API standards and found out that the API predictions are slightly lower than the experimental data.

This brief review of the related literature indicates the shortage of information in the field of storage tank evaporative losses. Furthermore, the above mentioned studies are mainly focused on the general estimation of the losses. However, in the present paper a numerical method has been developed for solving the energy equation to predict the storage tank temperature and to estimate the evaporative losses. More importantly, the solar irradiation and the effects of the tank surface paint absorptivity on the tank temperature and the evaporative losses have been investigated.

II. THE CASE STUDY

The problem under consideration is a typical storage tank in Khark Island shown in Figure 1. As seen, the exterior surface paint of the tank is white with two small rings of blue and yellow color indicating that the tank is suitable for storing both heavy and light crude oil.



Figure 1- The oil storage tank under consideration in Khark Island

The tank is an external pontoon floating roof type with 114 meters in diameter and 17 meters in height with the storage capacity of 1 million barrels of crude oil. The types and the numbers of deck fittings are listed in table 1.

Fitting Type	Construction Details		Number
Access hatch	Bolted cover, gasketed		3
Vacuum breaker	Weighted mechanical actuation, gasketed		14
Roof drain	100% open		5
Unslotted Guide Pole	Ungasketed sliding cover		2
Deck leg	Adjustable, pontoon area - gasketed		301
Rim vent	Weighted mechanical actuation, gasketed		13
Rim-seal	Primary	Liquid-mounted seal	1
	Secondary	Weather shield	

Table 1- fitting types of the tank



Figure 2- The deck leg of tank



Figure 3- The vacuum breaker of tank

Two fitting types of the tank are shown in Figures 2 and 3. Fig 2 show deck leg of tank, the exiting of crude oil vapor from gasketed area case to blacked the near area of gasketed, figure also show exposed liquid on the tank internal walls that vaporize as time goes on. In Fig 3 show vacuum breaker. The evaporation loss from this part lead to dirty around it, also the exiting of vapor could see in shadow of vacuum breaker. During the current study light crude oil with API of 33.36 has been stored in the tank.

III. THE NUMERICAL METHOD

The solar radiation is the main cause of the evaporative losses in the floating roof tanks. Theoretically, the solar radiation striking the earth atmosphere brings about 1.5 kW per square meter, when measured normal to the sun rays. This incident radiation is partly reflected and scattered and partly absorbed by the atmosphere [11]. For estimating solar radiation on the earth surface several engineering models have been proposed [12]. In all of the models the weather condition and geographic location are important factors. Kamali and Moradi [13] have examined various models including Angstrom, Bristow and Campbell, Hargreaves and Reddy for locations and weather conditions relevant to the present problem and compared their finding with the experimental data. It was suggested that Angstrom model with some modifications is more suitable for Khark Island conditions, and thus has also been adopted for the present study.

Based on the Angstrom model, solar radiation, H , can be estimated using the following equation:

$$\frac{H}{H_o} = a + b \frac{S}{S_o} \tag{1}$$

Where a and b are coefficients that must be chosen according to the location and weather conditions, S and S_o , are average sunshine duration and cloudless sunshine duration, respectively. Following Kamali and Moradi [13], a and b for Khark Island shown in table 3.

Table 2 – Coefficients that adjusted for Khark Island from Angstrom model

coefficient	spring	summer	autumn	winter
a	0.37	0.37	0.37	0.37
b	0.35	0.35	0.38	0.38

The cloudless hourly global irradiation received can be calculated using the following equation:

$$H_o = \frac{24 \times 3600}{\pi} \cdot I_{sc} \left(\frac{\bar{d}}{d} \right)^2 \left[\cos \phi \cdot \cos \delta \cdot (\sin \omega_2 - \sin \omega_1) + \frac{2\pi(\omega_2 - \omega_1)}{360} \cdot \sin \phi \cdot \sin \delta \right] \tag{2}$$

Where I_{sc} is set to 1367 W/m² according to the world radiation center [13] and ω is given by the following equation:

$$\omega = (t - 12) \times 15 \tag{3}$$

Figs. 4 and 5 show the time variations of the solar radiation, H , throughout the 5th day of each month in spring-summer and autumn-winter months throughout the year 2010, respectively. The solar radiation usually exists between 5:30 am to 18:30 pm with the pick around noon. Figures also show that the largest solar radiation occurs in June.

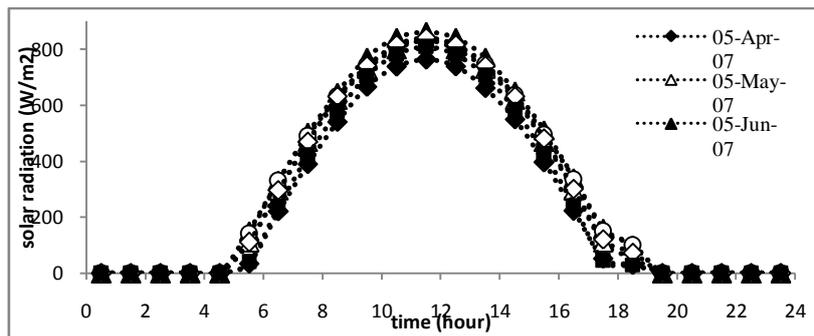


Figure 4 - Time variations of solar radiation in spring and summer months for Khark Island

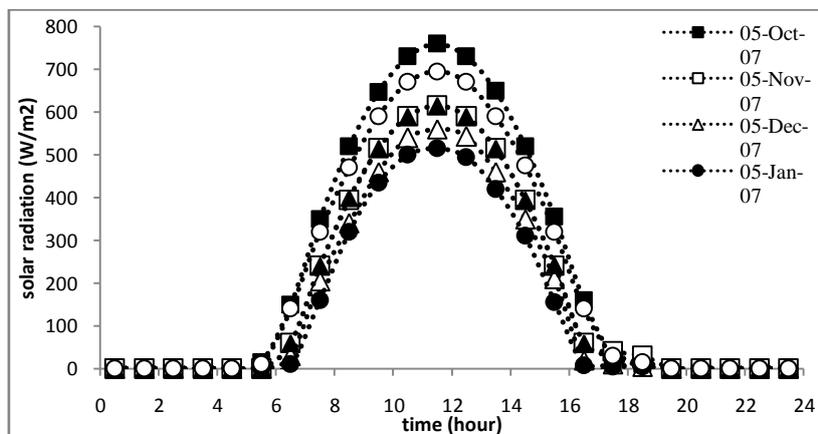


Figure 5- Time variations of solar radiation in autumn and winter months for Khark Island

A schematic diagram of the crude oil storage tank with all incoming and outgoing forms of energy is shown in Fig. 6. In developing the energy balance of the tank, the oil temperature variation inside the tank is neglected and a lumped system with uniform temperature is considered.

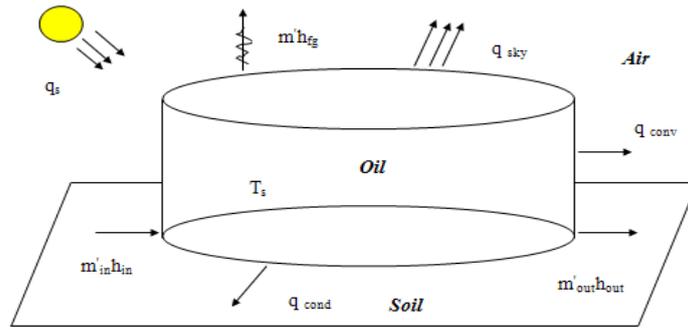


Figure 6 - Schematic of oil storage tank with all incoming and outgoing energies

Considering the tank as an open system, the energy equation can be expressed as:

$$\dot{Q} - \dot{W} + \dot{m}_{in} h_{in} - \dot{m}_{out} h_{out} - \dot{m} h_{fg} = \frac{dU}{dt} \quad (4)$$

Where \dot{Q} includes all incoming and outgoing heat fluxes expressed as:

$$\dot{Q} = q_s - q_{cond} - q_{conv} - q_{sky} \quad (5)$$

Where q_s is the absorbed solar energy by the tank surface with absorption coefficient, α , and irradiating surface area of A_s , with solar radiation, H , defined as:

$$q_s = \alpha \times A_s \times H \quad (6)$$

q_{cond} is the amount of heat conducted to the foundation ground evaluated by Fourier's law of heat conduction:

$$q_{cond} = -k A_b \frac{dT}{dx} = k A_b \frac{T_{soil} - T_s}{\Delta x} \quad (7)$$

Where Δx , k , and T_{soil} are thickness, conductivity coefficient and temperature of foundation base with the area of A_b , respectively.

q_{conv} evaluates the convective exchange of energy between the tank and the ambient with temperature T_∞ :

$$q_{conv} = h A_s (T_s - T_\infty) \quad (8)$$

There are many correlations available for calculating the convective heat transfer coefficient, h , in the above equation. In this study the correlation proposed by Churchill and Bernstein [14] has been employed, which is valid for vertical cylinders, when $RePr > 0.2$ related to the present case and expressed as:

$$\bar{Nu}_D = 0.3 + \frac{0.62 Re_D^{0.5} \cdot Pr^{1/3}}{\left[1 + \left(\frac{0.4}{Pr}\right)^{2/3}\right]^{1/4}} \cdot \left[1 + \left(\frac{Re_D}{282000}\right)^{5/8}\right]^{4/5} \quad (9)$$

Radiation heat exchange between the sky and the tank can be obtained according to:

$$q_{sky} = \sigma \cdot \epsilon \cdot A_s \cdot (T_s^4 - T_{sky}^4) \quad (10)$$

Where T_{sky} is the sky temperature evaluated following Kamali and Moradi [13] as:

$$T_{sky} = 0.0552 \times T_\infty^{1.5} \quad (11)$$

It is further assumed that the inflow and outflow rates of the crude oil are almost equal and therefore, the work done by the ambient pressure due to the negligible displacement of the tank roof related to the evaporative losses is neglected. Furthermore, the inflow enthalpy is assumed almost equal to the outflow enthalpy due to the small temperature differences. This assumption is also supported by the three dimensional numerical flow simulations inside the tank, which indicate that a large portion of the mass inflow to the tank directly moves toward the outflow region and does not mix considerably with the stored liquid, therefore:

$$\dot{W} = 0 \quad \text{and} \quad \dot{m}_{in} \dot{h}_{in} = \dot{m}_{out} \dot{h}_{out} \quad (12)$$

The time variation of the internal energy of the tank is expressed as:

$$\frac{dU}{dt} = \frac{d}{dt} (m \cdot c_p T_s) = \dot{m} \cdot c_p T_s + m \cdot c_p \cdot \frac{dT_s}{dt} \quad (13)$$

For simplicity, the quasi steady state condition has been assumed for the temperature time variation, which leads to:

$$\frac{dT_s}{dt} = 0 \quad \rightarrow \quad \frac{d}{dt} (m \cdot c_p T_s) = \dot{m} \cdot c_p T_s \quad (14)$$

The specific heat, c_p , of the crude oil is assumed to be a function of temperature as will be discussed later. The final form of the energy balance is obtained by substituting the above mentioned relations for each term in the energy equation (4) as follow:

$$\alpha HA_S + kA_b \frac{T_{soil} - T_S}{\Delta x} - hA_S(T_S - T_\infty) - \alpha \varepsilon A_s (T_S^4 - T_{sky}^4) - \dot{m} h_{fg} = \dot{m} c_p T_S \quad (15)$$

Based on the crude oil chemical composition as listed in Table 2 and with the use of HYSYS software, the following relations have been developed for the temperature dependence of evaporation enthalpy and specific heat:

$$h_{fg} = 1910T_S - 2773000 \quad (16)$$

$$c_p = 4.348T_S - 635.2 \quad (17)$$

Assuming $T_{soil} = T_\infty$ and incorporating h_{fg} and c_p in the energy equation yields:

$$\alpha \varepsilon A_s T_S^4 + 4.348 \dot{m} T_S^2 + (1910 \dot{m} + hA_S + \frac{kA_b}{\Delta x} + 635.2 \dot{m}) \times T_S - (\alpha HA_S + kA_b \frac{T_\infty}{\Delta x} + hA_S T_\infty + \alpha \varepsilon A_s T_{sky}^4 + 2773000 \dot{m}) = 0 \quad (18)$$

The storage tank temperature, T_S , and the crude oil evaporation rate, \dot{m} , are the two unknowns of the equation. Therefore, another equation is required to close up the problem. An equation will be developed for the evaporation rate following the API method as will be discussed in the next section. Evaporative losses from the external floating roof design are limited to losses from the seal system and roof fittings (standing storage loss) and any exposed liquid on the tank walls (withdrawal loss) [4, 5]. According to the API standards [15, 16], the total rates of evaporative losses from external floating roof tanks are equal to the sum of the rim seal losses, withdrawal losses, and deck fitting losses:

$$\dot{m} = \dot{m}_R + \dot{m}_{WD} + \dot{m}_F \quad (19)$$

Rim seal loss from floating roof tanks can be estimated using the following equation:

$$\dot{m}_R = 2.5 \times 10^{-8} \times (k_{Ra} + k_{Rb} V^n) DP^* M_V K_C \quad (20)$$

Where K_C is a product factor and for crude oil is set to 0.4, while K_{Ra} , K_{Rb} , and n are related to the used seal type. The vapor pressure, P^* , is evaluated according to:

$$P^* = \frac{\frac{P_{VA}}{P_A}}{\left[1 + \left(1 - \frac{P_{VA}}{P_A} \right)^{0.5} \right]^2} \quad (21)$$

Where the true vapor pressure, P_{VA} , for selected petroleum at the stored liquid surface temperature can be determined using the following equation:

$$P_{VA} = 6.895 \times \exp \left[A - \left(\frac{B}{T_S} \right) \right] \quad (22)$$

The constants A and B can be calculated from the following equations:

$$\begin{cases} A = 12.82 - 0.9672 \ln(RVP) \\ B = 7261 - 1216 \times \ln(RVP) \end{cases} \quad (23)$$

Where RVP is crude oil property, which is determined experimentally (given in Table 2). Deck fitting losses from floating roof tanks can be estimated according to:

$$\dot{m}_R = 2.5 \times 10^{-8} F_F P^* M_V K_C \quad (24)$$

The value of F_F is calculated using the actual tank-specific data for the number of each fit type (N_{Fi}) multiplying by the fitting loss factor for each fitting (K_{Fi}).

$$F_F = [(N_{F1} K_{F1}) + (N_{F2} K_{F2}) + \dots + (N_{Fn} K_{Fn})] \quad (25)$$

The deck fitting loss factor, K_{Fi} for a particular type of fitting, can be estimated by the following equation:

$$K_{Fi} = K_{Fai} + K_{Fbi} (K_V V)^{mi} \quad (26)$$

For external floating roof tanks, the fitting wind speed correction factor, K_v , is equal to 0.7. The withdrawal losses from floating roof storage tanks can be estimated using the following:

$$\dot{m}_{WD} = 2.829 \times 10^{-7} \times \frac{QC_s W_L}{D} \left[1 + \frac{N_C F_C}{D} \right] \quad (27)$$

Where N_C is zero for the external floating roof.

Finally, total rate of evaporative losses from external floating roof tanks can be calculated as:

$$\dot{m} = 2.5 \times 10^{-8} [F_F + D(k_{Ra} + k_{Rb} V^n)] P^* M_v K_C + 2.829 \times 10^{-7} \frac{Q C_s W_L}{D} \quad (28)$$

The vapor pressure function can be simplified by combining Equation (21) with Equation (22). The result is following equation:

$$P^* = 4 \times 10^{-7} \times e^{0.043 T_s} \quad (29)$$

Therefore, the total rate of evaporation losses can be estimated according to:

$$\dot{m} = 2.829 \times 10^{-7} \frac{Q C_s W_L}{D} + [F_F + D(k_{Ra} + k_{Rb} V^n)] M_v K_C \times 10^{-14} \times e^{0.043 T_s} \quad (30)$$

In the above equation, the evaporative rate, \dot{m} , is given as a function of the tank temperature. All other terms are constant coefficients which are determined based on the problem specifications according to the API method.

Equation 30 along with the energy balance, equation 18, is adequate to determine the tank temperature and the evaporative losses. Yet, an iterative method of trial and error is required to solve the set of equations. Consistent with the current color of the tank paint, the absorptivity of the tank is assumed to be 0.1 and 0.9 in calculation respectively.

It worth mentioning that in API method [15, 16], a simple correlation has been proposed for computing monthly averaged tank temperature, which can also be used for estimating the monthly averaged evaporative losses:

$$T_s = T_\infty + 1.86\alpha + 5 \times 10^{-5} \alpha H - 0.31 \quad (31)$$

IV. RESULTS AND DISCUSSION

As the first step in the evaluation of the evaporative losses, it is required to examine the accuracy of the developed numerical scheme for the tank temperature estimation. For this matter, the calculated monthly based averaged tank temperatures have been compared with their corresponding values from the API correlation given by equation 31, in Figure 7. Considering the approximate nature of both methods the results compare fairly well with each other. It must be emphasized that it is only the numerical method that provides a proper base for the study of the absorptivity of the paint effects on the transient tank temperature.

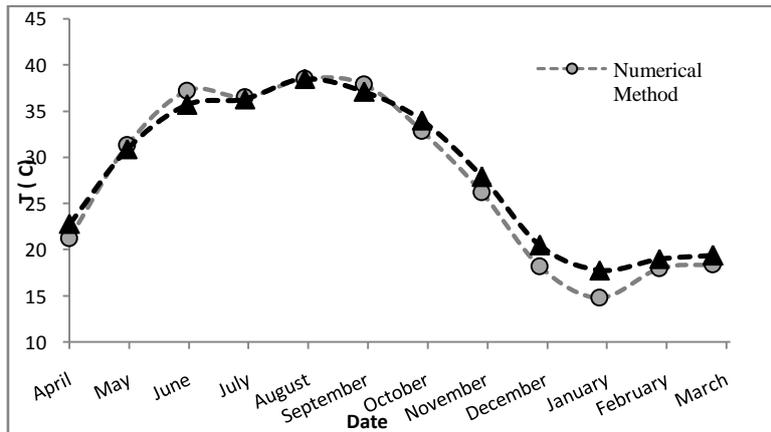


Figure 7 – Comparison between API and numerical method of monthly average tank temperature

Employing the numerical method discussed earlier the hourly variations of tank temperature have been plotted in Figs. 8 and 9 just for the 5th day of the spring and summer months, respectively. The corresponding ambient temperature variations have also been presented by dashed lines for comparison. Clearly, the tank temperature at the early and the late hours of the day is lower than the ambient temperature, which can be attributed to the exchange of radiative heat between the tank and sky with much lower value than the ambient temperature. As the sun rises, the solar radiation increases and so does the absorbed heat by the tank surface, which leads to the temperature rise of the tank. It must be emphasized that the tank temperature is also influenced by the wind speed, which can even outweigh the solar radiation. This fact is clear from Fig. 12, which shows that the highest tank temperature occurs at about noon in September, while the solar radiation is largest in June. This is due to fact that wind speed is much lower in September as compared to June.

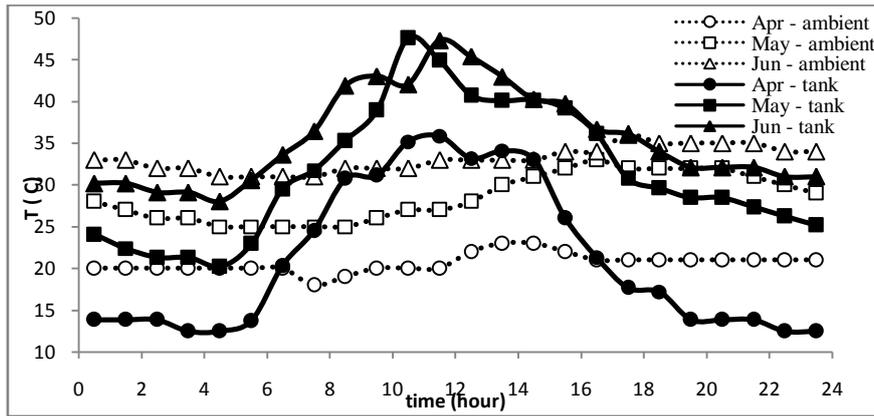


Figure 8 – Variation of the tank and ambient temperature against time in spring

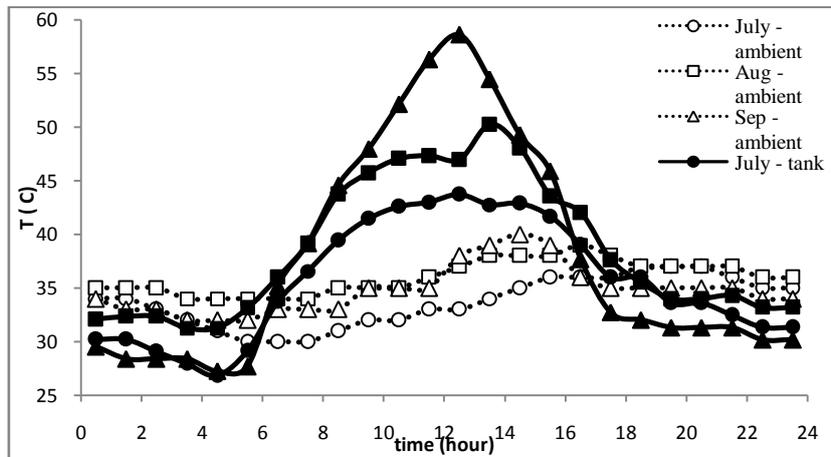


Figure 9 - Variation of the tank and ambient temperature against time in summer

Trough the developed numerical scheme the effects of the surface tank paint absorptivity on the tank temperature can be investigated, which enables the designer to select the appropriate paint that best suits the desired application. Fig. 10 shows the effects of paint absorptivity on time variations of the tank temperature during the 5th day of August 2010. It can be realized that absorptivity has strong effects on the tank temperature. The tank temperature can increase to about 40°C above the ambient temperature around 13:30 pm, when absorptivity increases from 0.1 to 0.9. Fig. 11 also shows the effects of the paint absorptivity but on the monthly averaged tank temperature. Similar to the time variation of the tank temperature the monthly averaged tank temperature can be increased by about 35°C in September, when absorptivity increases from 0.1 to 0.9.

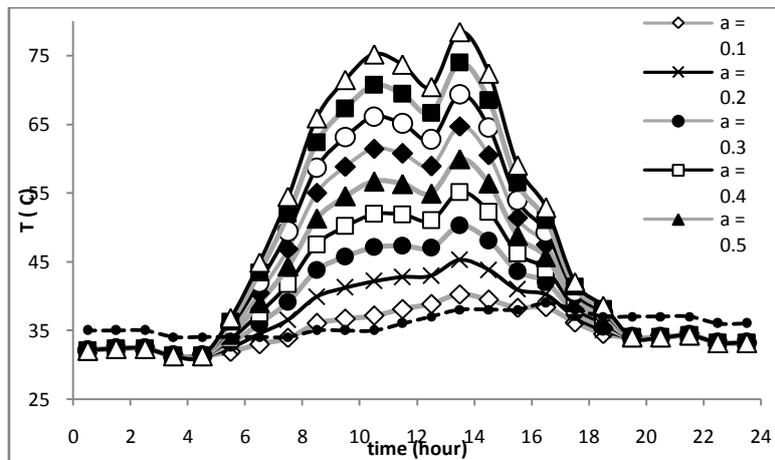


Figure 10 –Effects of absorptivity on time variations of the storage tank temperature for August 5th 2010

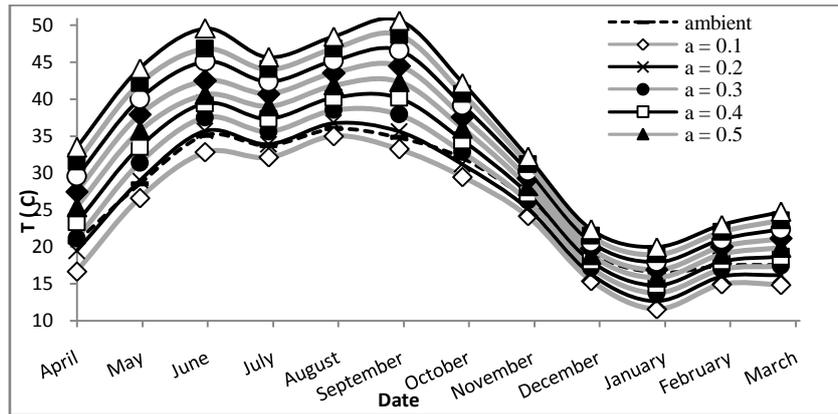


Figure 11 – Surface paint absorptivity effects on the variations of the monthly averaged tank temperature

Having examined the storage tank temperature, the evaporative losses from storage tanks can now be determined by the numerical method discussed earlier. In Fig. 12 the time variations of the evaporative losses during the 5 August 2010 are presented. The effects of the surface paint absorptivity of the storage tank on the evaporative losses can also be examined by the developed numerical method. Fig. 12 shows the effect of absorptivity of the storage tank paint on the evaporation rate from the tank during the 5th day of August 2010. It is clear that absorptivity dramatically affects the evaporation rate. An increase of about 0.18 bbl is observed, when absorptivity of the paint increases from 0.1 to 0.9.

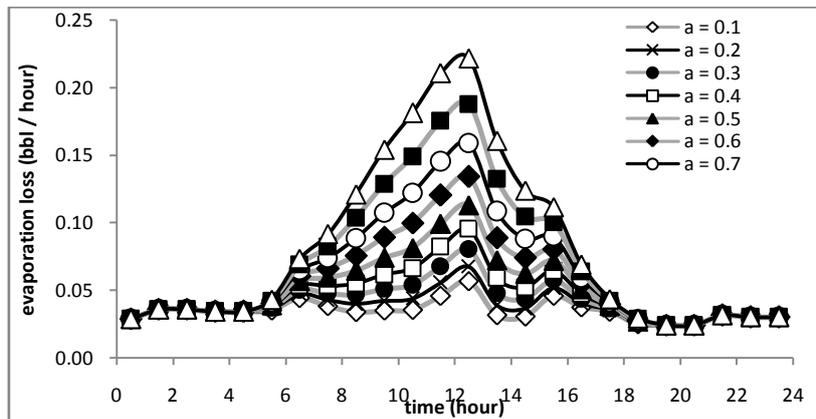


Figure 12 – Surface paint absorptivity effects on time variations of the storage tank temperature for August 5th 2010

Fig. 13 shows the variations of the monthly averaged evaporative losses from the storage tank throughout the year 2010. A comparison has also been made with the results obtained from the API method. It is expected to see that the highest evaporative losses are occurred during June, July and August, the hottest months in Khark Island. However, the local peaks in evaporative losses during October and November are due to the high wind speeds in these months. Furthermore, reasonable agreements between the results of the two methods are observed.

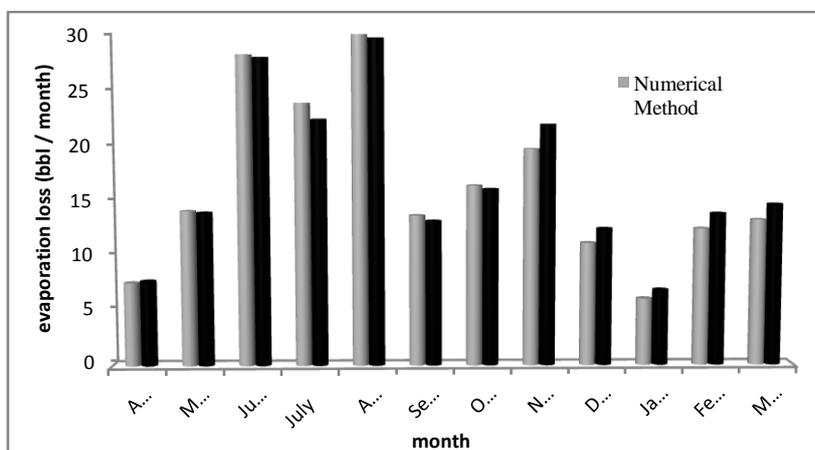


Figure 13 –Comparison of the monthly averaged total evaporative losses

Fig. 14 shows the effects of the paint absorptivity of the storage tank on the monthly averaged evaporative losses throughout the year 2010. As figure indicates, the absorptivity strongly influences the evaporation rate especially during the summer months. The evaporative losses increases by about 300%, when the absorptivity of the paint increases from $\alpha = 0.1$ to $\alpha = 0.9$ in August. Note that the evaporative losses in February and December are the same for all absorption coefficients due to the similar solar radiation and wind condition in these months. This is almost the case for the months of July and June.

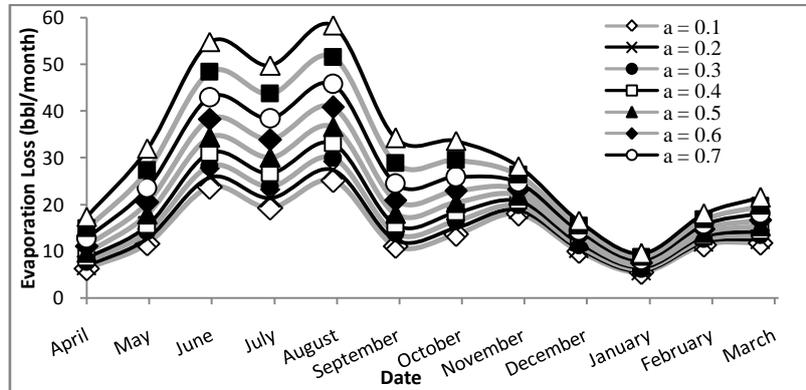


Figure 14 – Surface paint absorptivity effects on monthly variations of the evaporative losses

Annual averaged evaporative losses from the storage tank for different absorption coefficients are shown in Figure 15. A curve is fitted to the numerical values, which is expressed with the following simple expression.

$$Loss(barrel) = 200.7\alpha^2 + 51.59\alpha + 161.9 \quad (32)$$

The results indicate that simply by painting exterior surfaces by clean white ($\alpha=0.2$), the annual evaporative losses reduce by about 20 barrels as compared to the dirty white ($\alpha=0.3$). There are about 200 barrels difference between the light color ($\alpha=0.1$) and the dark color ($\alpha=0.9$) paints for a single storage tank. Since there are about 40 storage tanks in the Khark island total evaporative losses can be considerable. Furthermore, the amount of evaporative losses for lighter hydrocarbons is probably much higher than crude oil which calculated here, therefore, the absorptivity of the exterior surface paint of the storage tanks for such products plays more important roles, which must be considered.

Considering the fact that the evaporation basically occurs in the crude oil layers adjacent to the floating roof, the local temperature in this area is a key factor in this process. Clearly, the insulation of the tank roof, which absorbs the largest portion of the solar irradiation, reduces its temperature and the amount of evaporative losses accordingly.

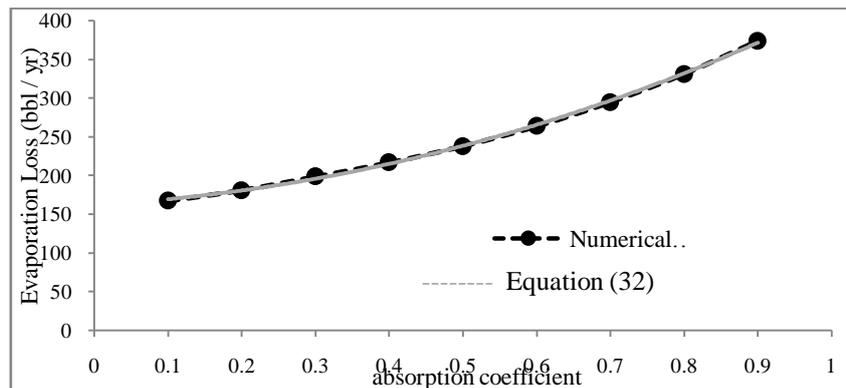


Figure 15- Annual evaporative losses as a function of the surface paint absorption coefficient

V. CONCLUSIONS

In this study a numerical scheme has been developed for estimating the time variations of the storage tank temperature and evaporative losses. The numerical value of monthly averaged evaporation losses have been compared with the estimations based on the API AP-42 standard. Considering the fact that the evaporation mainly occurs at the fluid surface under floating roof, therefore, surface temperature becomes important parameter. Any mechanisms that reduce the floating roof temperature will directly affect the evaporation rate. Therefore, the exterior surface paint absorptivity and even the cleaning of the floating roof, where dust can accumulate and increase the absorptivity of the surface becomes an issue. Present results indicate that the annual evaporative losses increase up to 125% if the absorptivity of the tank surface increases such that 90 percent of the solar irradiation is absorbed. It is expected that just by insulating the floating tank roof the evaporative losses reduce considerably. Furthermore, the evaporative losses are affected by the wind speed, which will reduce by adding wind shield system especially to the sealing system of the floating roof.

Nomenclature			
A_s	Area (m^2)	PA	atmospheric pressure, (kpa)
A	constant in the vapor pressure equation, (dimensionless)	P^*	vapor pressure function, (dimensionless)
B	constant in the vapor pressure equation ($^{\circ}K$)	Pr	Prandtl Number (dimensionless)
c_p	Special heat capacity (J/kg-K)	PVA	true vapor pressure, (kpa)
C_s	shell factor, (m)	Q	annual throughput, (m^3/yr)
D	tank diameter, (m)	q	Heat transfer energy (W/m^2)
F_c	effective column diameter, (m)	q_s	Absorbed solar energy (W/m^2)
F_F	total deck fitting loss factor, (kg-mole/yr)	q_{conv}	Convection Heat transfer energy (W/m^2)
H_0	cloudless daily global irradiation received, ($MJ/m^2 \cdot hour$)	q_{cond}	Conduction Heat transfer energy (W/m^2)
H	daily global irradiation, ($MJ/m^2 \cdot hour$)	q_{sky}	Radiation to sky (W/m^2)
h	convection coefficient ($w/m^2 k$)	Ra	Rayleigh number (dimensionless)
h_{fg}	evaporate enthalpy (KJ/kg)	S	average sunshine duration, (hour)
h_{in}	inlet enthalpy (KJ/kg)	S_0	cloudless sunshine duration, (hour)
h_{out}	Outlet enthalpy (KJ/kg)	T_{∞}	ambient temperature ($^{\circ}K$)
K	conductivity ($W/m K$)	TS	tank surface temperature, ($^{\circ}K$)
K_{Ra}	zero wind speed rim seal loss factor, (kg-mole/m@ yr)	T_{sky}	Sky temperature (K)
K_{Rb}	wind speed dependent rim seal loss factor, (kg-mole/(m/s) ⁿ m @yr)	T_{soil}	Soil temperature (K)
M_v	vapor molecular weight, (kg/kg-mole)	U	Internal energy (j)
m	Mass (kg)	V	average ambient wind speed, (m/s)
\dot{m}	total loss, (kg/s)	\dot{W}	Rate of work (W)
m_F	deck fitting loss, (kg/s)	WL	average organic liquid density, (kg/m^3)
m_R	rim seal loss, (kg/s)	Greek Symbols	
m_{WD}	withdrawal loss, (kg/s)	α	Absorption coefficient, (dimensionless)
\dot{m}_{in}	Inlet mass (kg/s)	δ	declination angle, (degree)
\dot{m}_{out}	Outlet mass (kg/s)	φ	Longitude, (degree)
N	Number of day in year	ω	hour angle, (degree)
NC	number of fixed roof support columns, (dimensionless)	ϵ	emissivity (dimensionless)
Nu	Nusselt number (dimensionless)	σ	Stefan–Boltzmann constant ($W/m^2 K^4$)

Reference

- [1] H. K. Abdel-Aal, M. Aggour, M. A. Fahim, Petroleum and gas field processing, In: Chapter 8- Storage Tanks and Other Field Facilities, New York: Marcel Dekker Inc; p. 1 (2003).
- [2] R.J. Laverman, Emission Reduction Options for Floating Roof Tanks, Second International Symposium on Aboveground Storage Tanks, Houston; (1992).
- [3] S. Ramachandran, Reducing (controlling) vapor losses from storage tanks, 7th annual India Oil & Gas Review, Symposium & International Exhibition, Bombay, (IORS-2000).
- [4] Ciolek Michael, Emission Factor Documentation for AP-42 Section 7.1 Organic Liquid Storage Tanks, In: Chapter 2- Storage Tank Descriptions, U S Environmental Protection Agency; p.1-5 (2006).
- [5] S. Wongwises, I. Rattanaprayura, S. Chanchaona, An Evaluation of Evaporative Emissions of Gasoline from Storage Sites and Service Stations, Thailand, Thammasat Int. J Sc Tech. Vol 2, No 2, (July 1997).

- [6] H. Asharif, E. Zorgani, Adjustment of Process Variables to Reduce Evaporation Losses from High pour point-crude oil storage tanks, The 8th international conference on petroleum phase behavior and fouling, France, 10-14 (June 2007).
- [7] D. Digrado Brian, A. Thorp Gregory, The aboveground steel storage tank handbook, In: chapter 13: new field-erected aboveground storage tank products, New Jersey: John Wiley & Sons Inc; p. 127. (2004).
- [8] Evaporative Loss from External Floating Roof Tanks, Bulletin No. 2517, Third Edition, American Petroleum Institute, Washington, DC, (1989).
- [9] Evaporative Loss from Internal Floating Roof Tanks, API Publication 2519, American Petroleum Institute, Washington, DC, (March 1990).
- [10] S. Zareie, D. Mowla, J. Fathi, Practical Study of VOCs Emission from External Floating Roof Tanks, 5th international congress of chemical engineering, Shiraz, Iran; (2007).
- [11] Wang Xiaoxin, Kendrick Chris, Ogden Ray, Maxted James, Dynamic thermal simulation of a retail shed with solar reflective coatings. Thermal Engineering 1073-1066. (2008);
- [12] ekai, Sen, Solar Energy Fundamentals and Modeling Techniques, Londen: Springer Veriag. p. 47-140. (2008).
- [13] G. A. Kamali, E. Moradi, Solar radiation fundamentals and application in farms and new energy, Tehran: publication 21 century; (2005).
- [14] S. W. Churchill, M. A. Bernstein, Correlating Equation for Forced Convection from Gases and Liquids to a Circular Cylinder in Cross Flow, J Heat Transfer. 300-306. (1977);
- [15] Manual of Petroleum Measurement Standards, In: Chapter 19- Evaporative Loss Measurement, Section 2- Evaporative Loss From Floating Roof Tanks. Preliminary Draft, American Petroleum Institute, Washington DC; (1994).
- [16] Ciolek Michael, Emission Factor Documentation for AP-42 Section 7.1 Organic Liquid Storage Tanks. In: Chapter 3-Emission Estimation Procedure. U S Environmental Protection Agency; p. 9-11: 15-18. (2006).

Numerical Approach for Temperature Development of Horizontal Pipe Flow with Thermal Leakage to Ambient

Andisheh Tavakoli¹, Mohamadreza Baktash²

¹Department of Material Engineering, Amirkabir University of Technology, Tehran, Iran
Petropars Ltd. Company, Client of Oil and Gas Project

²Department of Mechanic Engineering, K.N.Toosi University of Technology, Tehran, Iran
Namvaran Company, Consultant of Oil and Gas Project

Abstract: In this work, the effect of pipeline surface paint color on temperature development of horizontal pipe flow is investigated numerically. The classical pipe flow problem with constant wall heat flux is presented in which thermal leakage to the ambient from the outer surface of the pipe through convection and radiation is taken into account. The effects emissivity and absorptivity of pipe surface have been studied. The temperatures of fluid and pipe surface in the pipe flow problem are calculated numerically and experimentally. The results show that the bulk temperature tends to a limiting value (i.e. does not rise linearly as in the classical constant heat flux problem). It is found emissivity and absorptivity of outlet pipe surface are predominant parameters in temperature development in pipe flow that can increase or decrease pipe surface and fluid temperature for subjected objective. Furthermore, variations of friction factor in pipeline which that exposed to solar radiation and wind stream are investigated. The results show that friction factor and pressure loss are mainly affected by absorptivity and emissivity of exterior surface of pipe flow.

Key words: solar radiation, temperature development, pipe flow, emissivity and absorptivity, friction factor, pressure drop

I. Introduction

The pipe flow is a classical heat transfer problem and has been studied in all heat transfer textbooks (e.g. Incropera and DeWitt (2002) & Kays and Crawford (1993)). Generally, two thermal boundary conditions are applied for the pipe surface: constant wall heat flux and constant wall temperature. In either case, analytical solutions have been derived both for determination of the convective coefficient of hydrodynamically and thermally fully developed laminar flow, as well as for the temperature distribution along the pipe axis. The constant heat flux condition can be produced by a heater element or radiation over the tube surface. For the latter case, thermal leakage to the surroundings cannot be neglected.

The influence of solar radiation on pipe flow is widely applied to solar devices designs and analyses, especially in cylindrical solar collector and much work has been carried out in this field. For instance, Kocifaj (2009) submitted analytical solution for daylight transmission via hollow light pipes with a transparent glazing; analytical solutions was introduced for both transparent component and lambertian diffuser enables rapid numerical simulations for systems composed of several types of light guides. Bourdoukan et al. (2008) proposed a model for solar heat pipe vacuum collectors in the desiccant cooling process. Chirarattananon et al. (2000) suggested a model for receiving solar radiation by light pipe in tropics for increasing air-conditioning load. Madani (2006) calculated temperature distribution of water inside the cylindrical tube with black cover by simple model of overall coefficient of heat losses. Yaghoubi et al. (2003) estimated temperature of oil inside cylindrical receiver for Shiraz solar power plant; they supposed unsteady state condition and neglected conduction variation along pipe and obtained nonlinear partial differential equation system and solved by numerical method. Saroja et al. (1996) carried out unsteady analysis of a cylindrical solar water heater. The physical parameters that govern the physical system were identified. The governing equations were solved using the fourth order Runge-Kutta method for different values of the parameters. Kim et al. (2007) obtained thermal performance of a solar system composed of parallel, all-glass (double skin) vacuum tubes by using a one-dimensional analytical model; they supposed constant heat flux around tube under unsteady state condition; Han et al. (2008) continued same work with three dimensional model and compared results with one dimensional model and showed that there are good agreement with each other. Recently, temperature development in the pipe flow which there is dissipation of the heat through the outer surface to the ambient was studied by Kowsary and Pourshaghaghay (2007). They considered a more realistic situation in which there is dissipation of the heat (generated by the heater element) through the outer surface by using the first law and available thermal modeling tools. They found that bulk temperature tends to a limiting value along pipe.

In light of the aforementioned discussion, the objective of this research is to survey numerically for effect of solar radiation on temperature development of incompressible flow inside an aboveground pipeline. The outer surface of the pipeline is exposed to solar radiation and wind stream. The radiation heat exchange with ambient is also taken into account. The impacts of exterior surface paint color which represented by emissivity and absorptivity have been studied. For simplifying the model, hourly steady state situation in pipe flow has been assumed. In most studies such as Suehrcke et al. (2008) & Sahin and Kalyon (2005) because of using thermal resistance of radiation to sky, pipe surface temperature is assumed constant; while at this paper pipe surface temperature variation along flow is taken into account.

II. Analysis

A pipe flow which is exposed to ambient may gain heat through solar radiation of short wavelength over outer surface of the pipe during the day. The wind also plays a major role in opposite direction by removing heat from outer surface

through convection heat transfer. There is also exchange of heat between outer surface and ambient through long wavelength radiation heat transfer. Although the circular cross-section has been chosen here as the most common geometry in practice, the same analysis could be applied for other cross-sectional geometries. The inlet temperature of the fluid is T_{fi} and the outlet temperature is T_{fo} . Other parameters were shown as Fig. 1. It's considered that problem is three-dimensional at pipe and fluid flow.

Radiation heat flux can be expressed by considering the solar radiation of short wavelength and the ambient radiation of long wavelength and also radiation from ground as,

$$q_r = \alpha I_{solar} + \varepsilon \sigma (T_{sky}^4 - T_s^4) + I_{ground} \quad (1)$$

Where α and ε is the absorptivity and the emissivity of the surface paint color for radiation, while σ is the Stephan-Boltzmann constant and T_{sky} is the effective sky temperature that is given by Sharma and Mullick (1991) as,

$$T_{sky} = 0.0552 T_{\infty}^{1.5} \quad (2)$$

The solar radiation is the primary phenomena which effects temperature variation of surfaces. To be able to estimate pipe and fluid temperature, solar radiation should be calculated in first step. For estimating solar radiation many engineering models have been developed and proposed (Zekai (2008)). In all of these models, the weather condition and location are important factors. Iranian researchers have examined various models including Angstrom, Bristow & Campbell, Hargreaves and Reddy for Iran's cities and compared them with measured values (Kamali and Moradi (2005)). These models have been also modified by taking into account some other relevant meteorological variables and some suggesting coefficients for the difference regions. Kamali and Moradi (2005) suggested that Angstrom model is more suitable for the Kharg Island. Therefore, the Angstrom model has been employed in this study to calculate solar radiation. Based on Angstrom model, solar radiation can be estimated using the following equation as (Zekai (2008)),

$$\frac{I}{I_o} = a + b \frac{s}{s_o} \quad (3)$$

Where for Kharg Island, $a=0.37$ and $b=0.35$ in spring & summer, and $a=0.37$ and $b=0.38$ in autumn & winter (Kamali and Moradi (2005)).

The solar declination (δ), the main sunshine hour angle (ω_s) and the maximum possible sunshine duration day length (s_o) was calculated from Cooper (1969):

$$\delta = 23.45 \sin \left[\frac{360}{365} (284 + n) \right] \quad (4)$$

$$\omega_s = \text{Arc cos}(-\tan \phi \cdot \tan \delta) \quad (5)$$

$$s_o = \frac{2}{15} \omega_s \quad (6)$$

The cloudless hourly global irradiation received can be calculated using the following equation as (Duffie and Beckman (1991)),

$$I_o = \frac{12 \times 3600}{\pi} I_{sc} (1 + 0.033 \cos \frac{360n}{365}) [\cos \phi \cos \delta (\sin \omega_2 - \sin \omega_1) + \frac{2\pi(\omega_2 - \omega_1)}{360} \sin \phi \sin \delta] \quad (7)$$

In this equation I_{sc} has adopted a value of 1367 W/m^2 according to the world radiation center and ω is solar hour as given by following equation that recommended by Duffie and Beckman (1991) as,

$$\omega = 15(t - 12) \quad (8)$$

Solar energy can be varied by geographical situation. The Kharg Island condition is determined for numerical study, the pipe in the presented study is assumed to be located at Kharg island situation (Latitude: 29° and Longitude 51°). The model was solved for the climatic conditions of the Kharg Island, representing the southern Iran for the typical days (5th August 2008) of summer.

For calculation solar radiation on pipe surface following steps is taken into account. The total radiation on pipe surface consists of three components: beam solar radiation, sky diffuse radiation and ground reflected radiation. The total radiation calculated by:

$$I_T = I_b R_b + \frac{I_d}{2} + \rho_g \frac{I}{2} \quad (9)$$

While Reflectance of ground is considered as $\rho_g = 0.2$. The equations are presented in ASHRAE (1993), Handbook of Fundamentals. At above equation R_b is calculated as,

$$R_b = \frac{[1 - (\cos \phi \sin \delta - \sin \phi \cos \delta \cos \omega)^2]^{1/2}}{\cos \phi \cos \delta \cos \omega + \sin \phi \sin \delta} \quad (10)$$

The total hourly horizontal solar radiation, I , can be obtained from Eq. (3). However, data on beam solar radiation, I_b , and diffuse solar radiation, I_d , are not available. Therefore, a correlation between horizontal diffuse, I_d , and total horizontal solar radiation, I , is required. A correlation between horizontal diffuse and horizontal total solar radiation recommended by Boes (1979) is employed as follow:

$$\frac{I_d}{I} = 1 - 0.09K_T \quad \text{For} \quad K_T \leq 0.22 \quad (11)$$

$$\frac{I_d}{I} = 0.9511 - 0.1604K_T + 4.388K_T^2 - 16.638K_T^3 + 12.336K_T^4 \quad \text{For} \quad 0.22 < K_T \leq 0.8 \quad (12)$$

$$\frac{I_d}{I} = 0.165 \quad \text{For} \quad K_T > 0.8 \quad (13)$$

In order to estimate the convective heat transfer in outside pipe, the concept of mixed convection should be employed. The directions of air motion due to natural and forced convection are approximately perpendicular. The heat transfer coefficient from the pipe to the surrounding can be calculated by a correlation against wind speed as from Duffie and Beckman (1991):

$$h_o = 5.7 + 3.8V_w \quad (14)$$

The following assumptions for utilized model are considered:

1. The physical properties of all components of the pipe flow don't change with temperature.
2. Steady state situation assumed and heat transfer coefficients are considered to be constant at the selected time interval (hourly here).
3. The surface color is opaque with constant absorptivity and emissivity.
4. The variations in the absorptivity and emissivity of the color surfaces with the variation in angle of the incoming radiation are neglected.
6. The topper surface of pipeline is considered for numeric and experiment study.
5. Although emissivity and absorptivity are related to each other; but its effect is infinitesimal and so ignored.

Eqs. (15–19) are derived based on the above assumptions to examine the steady state behavior of pipe flow under investigation:

Continuity equation:

$$\frac{1}{r} \frac{\partial}{\partial r} (\rho r v_r) + \frac{1}{r} \frac{\partial}{\partial \theta} (\rho v_\theta) + \frac{\partial}{\partial z} (\rho v_z) = 0 \quad (15)$$

Momentum equation:

— r component:

$$\rho (v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} + v_z \frac{\partial v_r}{\partial z} - \frac{v_\theta^2}{r}) = \mu [\frac{\partial}{\partial r} (\frac{1}{r} \frac{\partial}{\partial r} (r v_r)) + \frac{1}{r^2} \frac{\partial^2 v_r}{\partial \theta^2} - \frac{2}{r^2} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial^2 v_r}{\partial z^2}] - \frac{\partial p}{\partial r} + \rho g_r \quad (16)$$

— θ component:

$$\rho (v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + v_z \frac{\partial v_\theta}{\partial z} - \frac{v_r v_\theta}{r}) = \mu [\frac{\partial}{\partial r} (\frac{1}{r} \frac{\partial}{\partial r} (r v_\theta)) + \frac{1}{r^2} \frac{\partial^2 v_\theta}{\partial \theta^2} + \frac{2}{r^2} \frac{\partial v_r}{\partial \theta} + \frac{\partial^2 v_\theta}{\partial z^2}] - \frac{1}{r} \frac{\partial p}{\partial \theta} + \rho g_\theta \quad (17)$$

— z component:

$$\rho (v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z}) = \mu [\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial v_z}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 v_z}{\partial \theta^2} + \frac{\partial^2 v_z}{\partial z^2}] - \frac{\partial p}{\partial z} + \rho g_z \quad (18)$$

Energy equation:

$$\rho (v_r \frac{\partial h}{\partial r} + \frac{v_\theta}{r} \frac{\partial h}{\partial \theta} + v_z \frac{\partial h}{\partial z}) = \alpha [\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial T}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 T}{\partial \theta^2} + \frac{\partial^2 T}{\partial z^2}] \quad (19)$$

To solve these rather complex equations, commercial computational fluid dynamics (CFD) package Fluent 6.0.12 (Ansys Inc.) based on the finite volume approach has been employed. Convection terms are approximated using upstream differences whereas a fully implicit scheme is introduced to handle transient terms. Also, pressure and velocity fields are linked together by the SIMPLE algorithm. Solutions are obtained for the system when it is assumed to have achieved a steady state operation with the lapse of time.

Due to obtain the pipe surface temperature profile and fluid inside, a numerical analysis has been employed. Fig. 2 shows a partial view of the three-dimensional grid system used in the present analysis. The system consists of 266,000 elements where non-uniform ones are employed in domains of high fluid motion and heat transfer. The generated mesh distribution was designed to give an optimal accuracy: small elements (2.4 mm) were used in the pipe, air and fluid, while larger elements (20 mm) were used for the terrain below the pipe flow. It is constructed according to the physical dimensions and properties of the pipe flow under investigation. In all cases considered in the present analysis, temperatures were initially assumed at 20 °C.

III. Results and discussion

In order to illustrate for the influence of absorptivity and emissivity variation and discuss the relevant parameters, a base case is selected for which the thermophysical parameters and pipeline dimensions are given in Table 1. The working fluid is chosen to be light oil. The thermal properties of oil at 15.5 °C are also brought in Table 1. The flow is assumed to be fully developed and laminar. For the base case the Reynolds number is calculated to be,

$$Re = \frac{\rho V D_i}{\mu} = \frac{858.4 \times 0.01 \times 0.0214}{0.0145} = 12.66$$

This ensures that the flow is laminar. The heat transfer coefficient on the inner surface of the pipe is then calculated to be,

$$h_i = \frac{k_{oil} Nu}{D_i} = \frac{0.1483 \times 4.364}{0.0214} = 30.24$$

The profile of the fluid temperature is affected by the selection of the outer convective and radiative heat transfer coefficients. Therefore, for the case study selected the radiation heat transfer is considered to be significant. (When studying absorptivity effect, ϵ set to 0.8 and for surveying the effect of emissivity coefficient, α set to 0.8). The paint color material thickness is very low in compare with pipe thickness such that when thermal resistance of color over the pipe is considered to be neglected.

The pipe flow system was evaluated for a day period on an hourly basis. The major output of the numerical analysis was the temperature profile for both the fluid and the pipe surface for the pipe flow. As an example, the steady state pipe temperature in the pipe flow as well as the ground temperature distribution which is located below the pipeline is shown in Figs. 4 during at 12:00 and 13:00 o'clock in 5th august 2008. At noon, solar radiation directly hit the upper surface of pipe and so no solar radiation will arrive to ground elements which are located below the pipe, hence the temperature of these elements will be lower than other elements. This matter is obvious in fig. 3 which can verify the numerical results anyhow.

Fig. 4 shows the variation bulk fluid temperature for the base case ($\epsilon=0.8$). As the inlet temperature of fluid entering the pipe is lower than the surrounding, the fluid bulk temperature along the pipe is expected to increase. It could be found that increasing absorptivity coefficient also growing bulk temperature along pipe. In addition, the bulk temperature inclined to a limiting value after some distance from the beginning of the pipe. Considering ambient temperature, it may be realized for high absorptivity coefficients (0.9, 0.7 ...), the bulk flow limiting temperature is much higher than ambient temperature. As for liquids, viscosity (and so pressure drop along the pipe) decreases as temperature increases, one way to reduce pumping work required for moving crude oil along long distance pipelines is to paint exterior surface with dark color.

Fig. 5 shows effect of absorptivity coefficients on variation of pipe surface temperature along flow for base case ($\epsilon=0.8$). As it is known, high absorptivity coefficient causes higher amount of solar radiation to be absorbed by pipe surface, therefore steady state pipe temperature at end of flow for $\alpha=0.9$ is higher than pipe temperature with lower absorptivity coefficients. Considering ambient temperature, it could be realized for low absorptivity coefficient (0.1), the pipe surface temperature is lower than ambient temperature for whole pipe length. The main reason for this behavior is due to exchange of heat (long wavelength radiation) between the pipe surface and sky which is at much lower than ambient temperature. This effect could be especially observed in clear sky (and night) in actual conditions.

Figs. 6 and 7 indicate effect of emissivity coefficients on variation of the bulk fluid and surface temperature for the base case ($\alpha=0.8$) respectively. As shown the effects of emissivity on temperature variation of fluid and pipe surface has inversely influence rather than absorptivity. Emissivity coefficient causes reflection of heat from pipe surface to sky, therefore as it increases, heat rejection from pipe surface to sky increases and so pipe surface and fluid temperature decreases. It could be found that lower emissivity coefficient is caused higher pipe surface temperature. Due to convey better insight for the problem under investigation, influences of regular paint color in industry on temperature development in the pipe flow are represented in Fig. 8. The corresponding absorptivity and emissivity for each color are listed in table 2. As expected, using black paint as envelop color of pipeline leads to intensify fluid flow temperature through pipeline.

In accordance with above figures, whatever proceed along the pipeline, it can be seen that the slope of fluid and pipe surface temperature is tended to straight line and at the end of pipe is converged to limited value. In fact, when there is heat dissipation through the radiation and convection to the ambient, the bulk temperature changes similar to exponentially, reaching a limiting value for large x (fig. 8-b). Moreover, both the fluid temperature and the tube surface temperature tend to an identical value far from the tube entrance. This is unlike the classical constant wall heat flux case in which temperature rises unboundedly in a linear manner (fig. 8-a). To check validity of the numerical results, a comparison has been made between the numerical results and measured values of outer pipeline surface temperature. The experimental temperatures have been measured using infrared thermometer from upper surface of pipeline on 5th august 2008 at 12:00 o'clock. The pipeline is considered to have off-white paint color accompanied by absorptivity and emissivity equals to 0.34 and 0.9 respectively. At that particular day, light crude oil flows through pipeline while the outlet surface temperature of pipe has

been measured along axial distance on the pipeline. The outlet surface temperature has been measured from pipeline entrance while its length is about 300 m. Comparison between the numerical and experimental results are depicted in Fig. 9. In this work, it is found that the results of the developed numerical model are in good agreement with the measured values.

The current section will try to discuss the effects of the pipe surface absorptivity and emissivity coefficients on transferring fluid through pipe in laminar flow using the previous described model. This may help engineers in their design to determine and select the optimum envelop color. The selection may be the prime interest of many engineering applications. Flow is always accompanied by friction. This friction results in a loss of energy available for work. A general equation for estimate pressure drop due to friction is the familiar Darcy equation as,

$$\Delta p = \frac{0.5 \rho f_m L V^2}{D_i} \quad (20)$$

That friction factor is obtained for laminar flow as below,

$$f_m = \frac{64}{\text{Re}} \quad (21)$$

Fluid temperature is an important parameter on variation of fluid viscosity and in consequence friction factor. The viscosity is decreased due to increasing temperature and so loss work could be diminished. Hence fluid flow pressure drop will be declined along pipe. Here, the effects of exterior surface emissivity and absorptivity upon fluid temperature are taken into account. As discussed previously, the fluid temperature tends to an identical value far from the tube entrance in pipe. Which this is depended on emissivity and absorptivity coefficients of surface paint color. Variation of viscosity and density versus temperature for the light oil is gained by following equations. The variation of properties by temperature is based on experimental analysis of the Iranian light crude oil.

$$\rho = -0.732T + 869.3 \quad (22)$$

$$\mu = 10^{-8}T^4 - 10^{-6}T^3 + 7 \times 10^{-5}T^2 - 0.002T + 0.033 \quad (23)$$

In current study, the temperature is assumed to be reached to its final value so viscosity and density could be assumed constant for each case. By knowing, viscosity and density, the friction factor, f_m , could be calculated by Eq. (21). Once the friction factor is calculated, the pressure drop for the studied pipeline could be easily obtained by Eq. (20). The influence of different absorptivity and emissivity coefficients of pipe exterior surface on the developed temperature, friction factor and eventually pressure drop are indicated in the table 3. It is seen that the pipe surface with higher absorptivity and lower emissivity bring about fluid temperature to increase; subsequently oil viscosity and loss work to decrease and therefore pressure loss will be decreased.

Considering the effects of absorptivity and emissivity of exterior surface which represents surface paint color on fluid bulk temperature development, it could be concluded that surface paint color should be selected based on the desired application. If the intention is to reduce fluid temperature, lighter color should be applied. For Kharg Island pipeline, any temperature rise of crude oil will have a negative impact on accuracy of flow measurement. So paint color with lowest absorptivity and highest emissivity is proposed. Like discussed, for oil pipeline the dark color should be applied to reduce pressure drop along the pipelines.

IV. Conclusions

In the present study, a numerical analysis for predicting temperature development for classical problem of fluid flow inside a pipe with constant wall heat flux by solar radiation in which there is leakage of the heat flux to the ambient has been established. The outer surface of the pipeline is exposed to solar radiation and wind stream. The radiation heat exchange with ambient is also taken into account. The effects of exterior surface paint color which represented by emissivity and absorptivity has been studied. The model has been developed to study crude oil flow temperature development along flow direction in an aboveground pipeline. The results of the numerical solution are in good agreement with those of the experimental model of the kharg oil pipeline with off white as envelope color.

The results obtained by the model show that the bulk temperature of the fluid and pipe surface along pipe varies exponentially in the direction of tube length. In the limit $x \rightarrow \infty$, the bulk temperature and surface temperature of the pipe do not increase unboundedly; rather, they tend to an asymptotic value. Based on the results which indicated significantly of exterior surface paint color, one should choose the paint color by considering its effects on temperature development. For increasing fluid temperature, the paint color that has high absorptivity and low emissivity should be used and for lower growing or constant temperature objective the paint color that has low absorptivity and high emissivity could be proposed.

Nomenclature

D_i	inner diameter of pipe (m)
D_o	outer diameter of pipe (m)
g	acceleration due to gravity (m/s^2)
f_m	Friction factor (dimensionless)
h	enthalpy (J/kg)
h_i	convection coefficient of inside pipe ($W/m^2 K$)
h_o	convection coefficient of air around pipe ($W/m^2 K$)
I	hourly solar radiation on horizontal surface (W/m^2)
I_b	beam solar radiation (W/m^2)
I_d	diffuse solar radiation (W/m^2)
I_o	extraterrestrial radiation (W/m^2)
I_{sc}	solar constant (W/m^2)
I_T	hourly solar radiation on pipe surface (W/m^2)
K_T	Daily average clearness index (dimensionless)
k_w	conductivity of pipe ($W/m.K$)
L	Length of pipeline (m)
Nu	Nusselt number ($= hD/k$)
n	number of day of the year, starting from the first of January
q_r	Radiative heat flux (W/m^2)
r, θ, z	Cylindrical coordinates
Re	Reynolds number, ($=VD/\nu$)
S	Daily average daily measured sunshine duration (h)
S_o	Daily average maximum possible sunshine duration (h)
T_f	Fluid temperature ($^{\circ}C$)
T_s	Pipe surface temperature ($^{\circ}C$)
V	Velocity (m/s)
V_w	Wind velocity (m/s)
P	Pressure (Pa)
ΔP	Pressure drop (Pa)

Greek symbols

ρ	density (Kg/m^3)
ϵ	emissivity (dimensionless)
σ	Stefan–Boltzmann constant ($5.67 \times 10^{-8} W/m^2 K^4$)
α	absorptivity (dimensionless)
ϕ	latitude of site ($^{\circ}$)
ω	solar angle ($^{\circ}$)
δ	solar declination angle ($^{\circ}$)
μ	dynamic viscosity, $kg/m.s$
ζ	z dimensionless axial distance (z/L)

Subscripts

r	radial
θ	angular
z	lengthwise

References

- [1] ASHRAE, (1993) *Handbook of Fundamentals*, Atlanta: American Society of Heating, Refrigeration and Air Conditioning Engineers Inc.
- [2] Boes, E.C. (1975) Estimating the Direct Component of Solar Radiation, *Sandia Report SAND.*, 75, pp. 55–65.
- [3] Bourdoukan, P. Wurtz, E. Joubert, P. and Sperandio M., (2008) Potential of solar heat pipe vacuum collectors in the desiccant cooling process: Modelling and experimental results, *Solar Energy.*, 82(2), pp. 1209–1219.
- [4] Chirarattananon, S. Chedsiri, S. and Renshen L., (2000) Daylighting through light pipes in the tropics, *Solar Energy.*, 69(3), pp. 331–341.

- [5] Cooper, P.I., (1969), The absorption of solar radiation in solar stills, *Solar Energy.*, 12, pp. 333–345.
- [6] Duffie, J.A. and Beckman, W.A., (1991) *Solar Engineering of Thermal Processes*, New York: Willy, p. 45–61.
- [7] Han, H. Kim, J.T. Ahn, H.T and Lee S.J., (2008) A three-dimensional performance analysis of all-glass vacuum tubes with coaxial fluid conduit, *International Communications in Heat and Mass Transfer.*, 35, pp. 589–596.
- [8] Incropera, F.P. and DeWitt, D.P., (2002) *Fundamentals of heat and mass transfer*, New York: John Wiley. p. 125–176.
- [9] Kamali, G.A. and Moradi, E., (2005) *Solar radiation fundamentals and application in farms and new energy*, Tehran: Teh, pp 136–146.
- [10] Kays, W.M. and Crawford, M.E., (1993) *Convective heat and mass transfer*, New York: McGraw Hill, p. 253–376.
- [11] Kim, J.T. Ahn, H.T. Han, H. Kim, H.T and Chun, W., (2007) The performance simulation of all-glass vacuum tubes with coaxial fluid conduit, *International Communications in Heat and Mass Transfer.*, 34, pp. 587–597.
- [12] Kocifaj, M., (2009) Analytical solution for daylight transmission via hollow light pipes with a transparent glazing, *Solar Energy.*, 83(2), pp. 186–192.
- [13] Kowsary, F. and Pourshaghagh, A., (2007) Temperature development in pipe flow with uniform surface heat flux condition considering thermal leakage to the ambient, *Energy Conversion and Management.*, 48, pp. 2382–2385.
- [14] Madani, H., (2006) The performance of a cylindrical solar water heater, *Renewable Energy.*, 31(2), pp. 1751–1763.
- [15] Sahin, A.Z. and Kalyon, M., (2005) Maintaining uniform surface temperature along pipes by insulation, *Energy.*, 30, pp. 637–647.
- [16] Saroja, S. Nithiarasu, P. and Seetharamu K.N., (1996) Transient analysis of a cylindrical solar water heater, *Energy Conversion and Management.*, 38(18), pp. 1833–1845.
- [17] Sharma, V.B. and Mullick, S.C., (1991) Estimation of heat-transfer coefficients, the upward heat flow, and evaporation in a solar still, *ASME Journal of Solar Engineering.*, 113, pp. 36–41.
- [18] Suehrcke, H. Peterson, E. and Selby, N., (2008) Effect of roof solar reflectance on the building heat gain in a hot climate, *Energy and Buildings.*, 40, pp. 2224–2235.
- [19] Yaghoubi, M. Azizian, K. and Kenary, A., (2003) Simulation of Shiraz solar power plant for optimal assessment, *Renewable Energy.*, 28, pp. 1985–1998.
- [20] Zekai, S., (2008) *Solar Energy Fundamentals and Modeling Techniques*, London: Springer, p. 180–248.

List of Tables

Table1. Parameters used in numerical study.

Table 2. Sample for various envelope color.

Table 3. Main results for the effect of absorptivity and emissivity on pressure drop and other related parameters.

List of Figures

Fig. 1. A schematic diagram of pipe flow under investigation.

Fig. 2. A section of mesh system for aboveground pipe flow.

Fig. 3. Temperature distribution in the pipe surface and ground during at 12:00 and 13:00 o'clock in 5th august 2008.

Fig. 4. Effect of absorptivity coefficient on temperature variation of bulk fluid along the pipeline.

Fig. 5. Effect of absorptivity coefficient on temperature variation of pipe surface along the pipeline.

Fig. 6. Effect of emissivity coefficient on temperature variation of bulk fluid along the pipeline.

Fig. 7. Effect emissivity coefficient on temperature variation of pipe surface along the pipeline.

Fig.8. Effect of pipe surface color on temperature variation of bulk fluid along the pipeline.

Fig. 9. temperature Variation of pipe surface and fluid along the pipe with constant wall heat flux, a) no heat dissipation (classical pipe flow problem), b) heat dissipation to ambient.

Fig. 10. Comparisons of the calculated and measured temperatures on the surface of the kharg oil pipeline in august 2008.

Table 1. Parameters used in numerical study.

Parameter	Value
L(m)	10
D _o (m)	0.0254
D _i (m)	0.0214
k _w (W/m.K)	30
T _∞ (°C)	36
I(W/m ²)	700

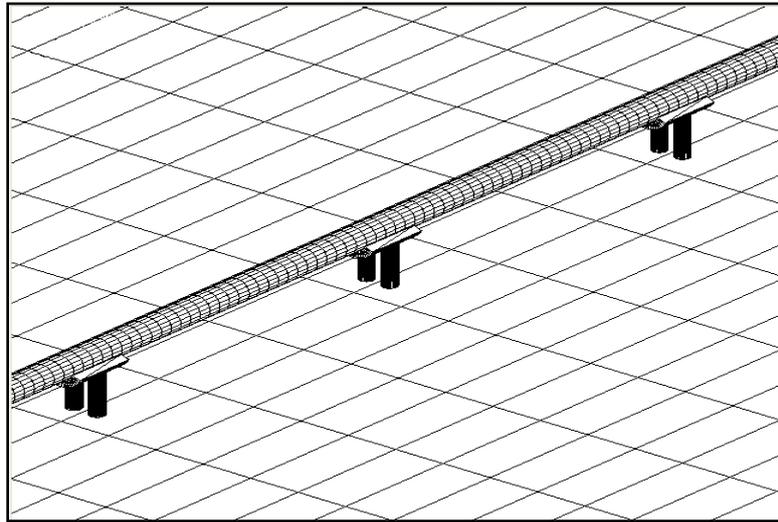


Fig. 2. A section of mesh system for aboveground pipe flow.

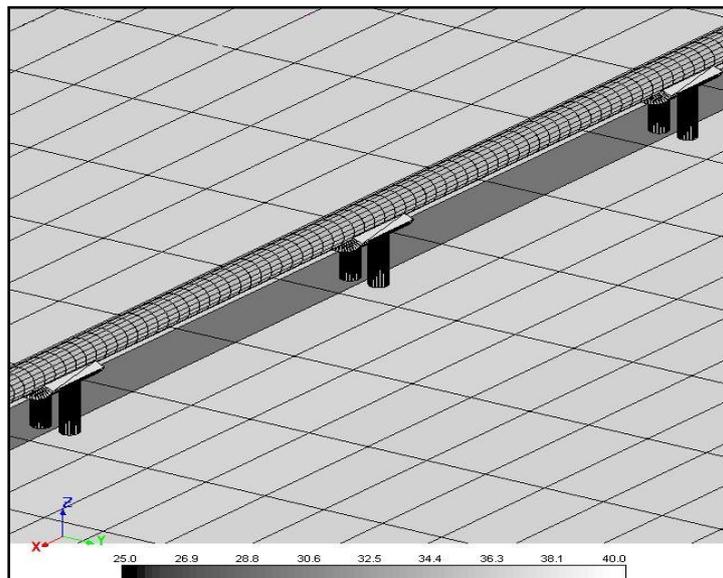


Fig. 3. Temperature distribution in the pipe surface and ground during at 12:00 and 13:00 o'clock in 5th august 2008.

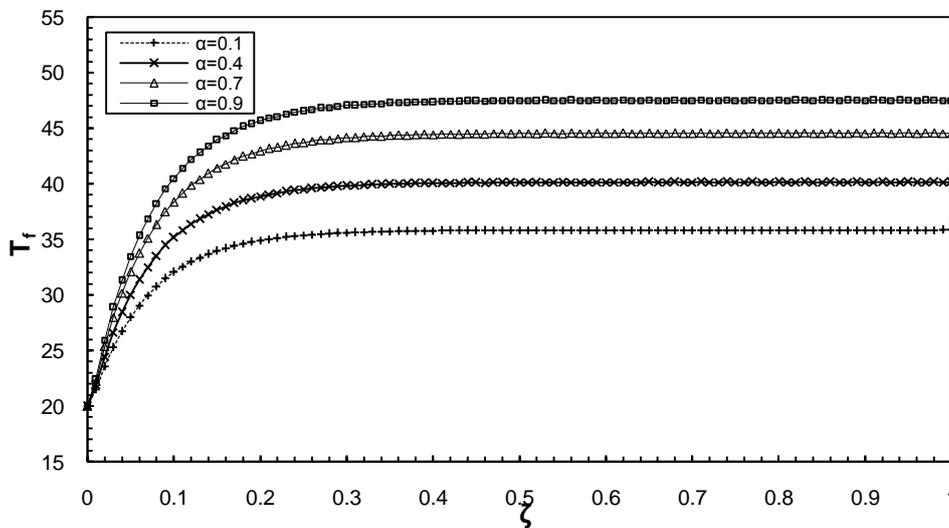


Fig. 4. Effect of absorptivity coefficient on temperature variation of bulk fluid along the pipeline.

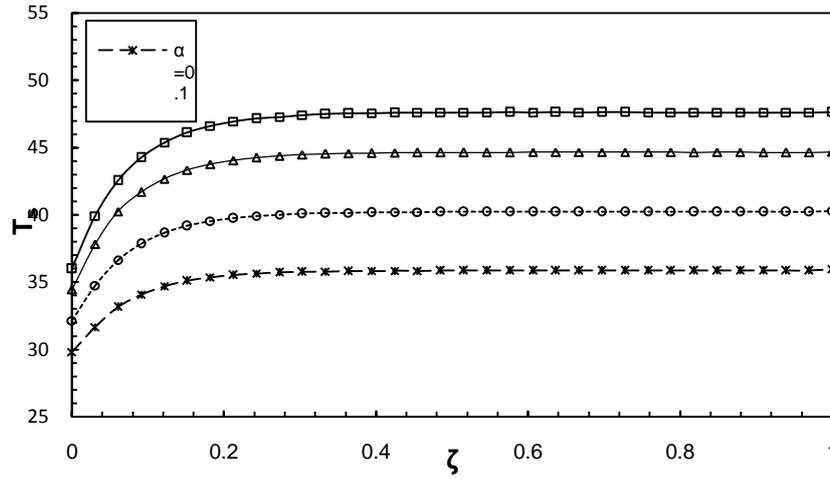


Fig. 5. Effect of absorptivity coefficient on temperature variation of pipe surface along the pipeline.

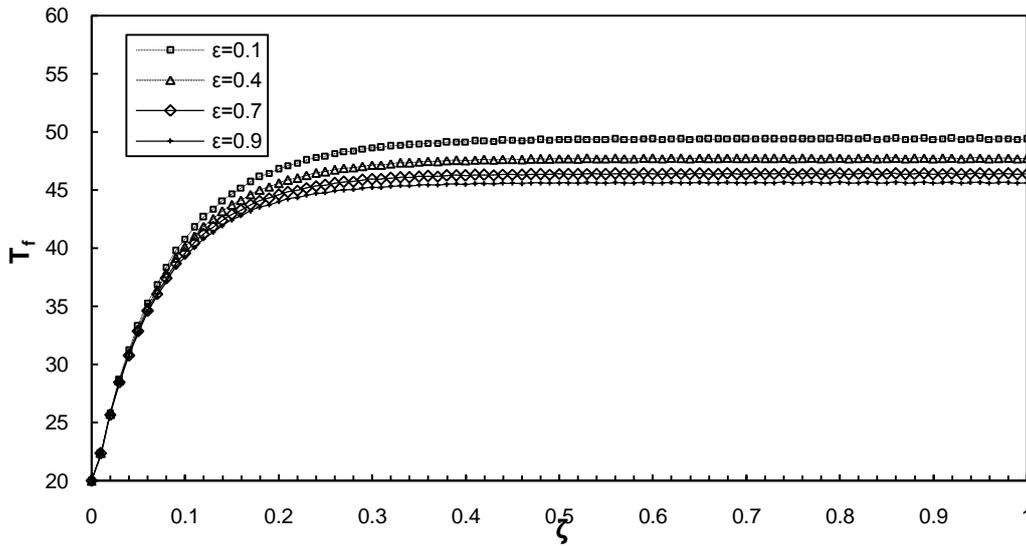


Fig. 6. Effect of emissivity coefficient on temperature variation of bulk fluid along the pipeline.

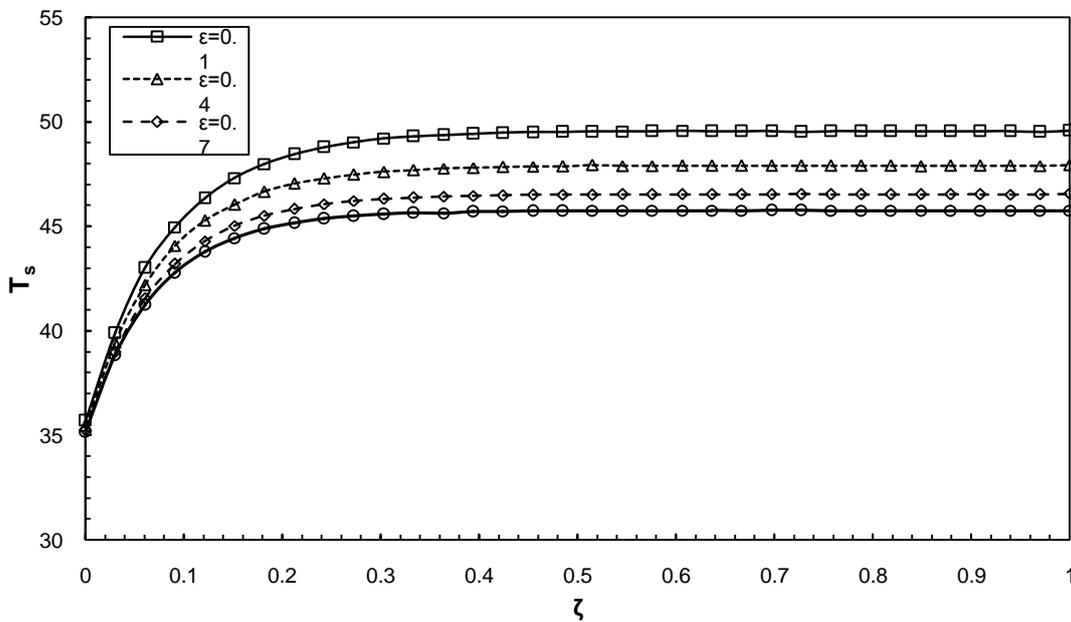


Fig. 7. Effect emissivity coefficient on temperature variation of pipe surface along the pipeline.

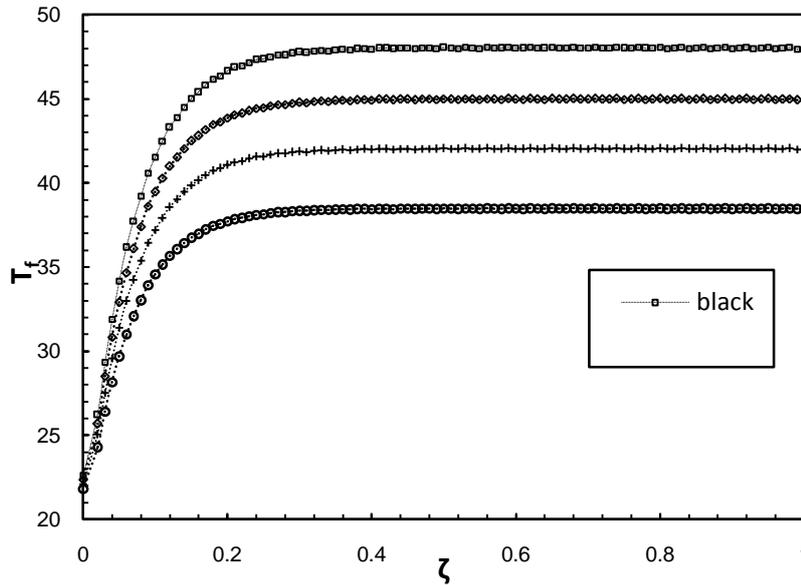


Fig. 8. Effect of pipe surface color on temperature variation of bulk fluid along the pipeline.

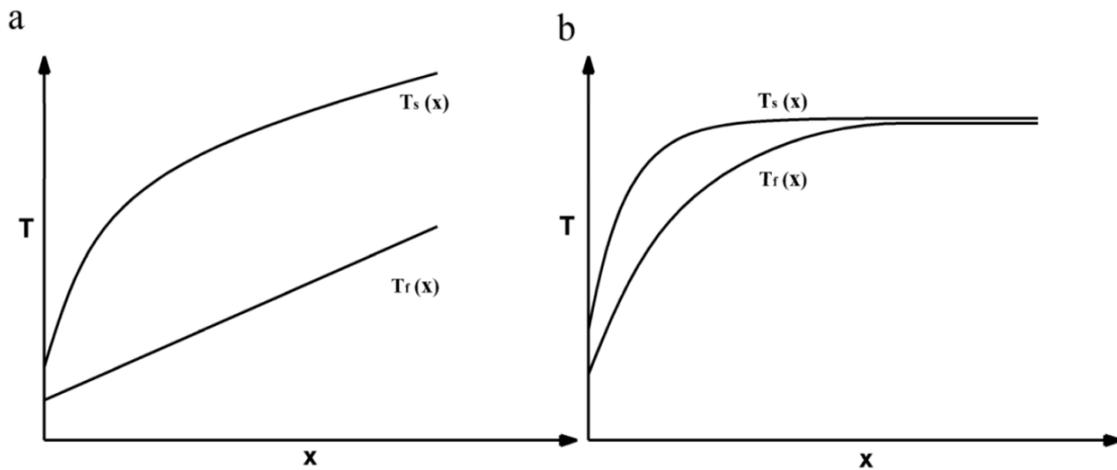


Fig. 9. temperature Variation of pipe surface and fluid along the pipe with constant wall heat flux, a) no heat dissipation (classical pipe flow problem), b) heat dissipation to ambient

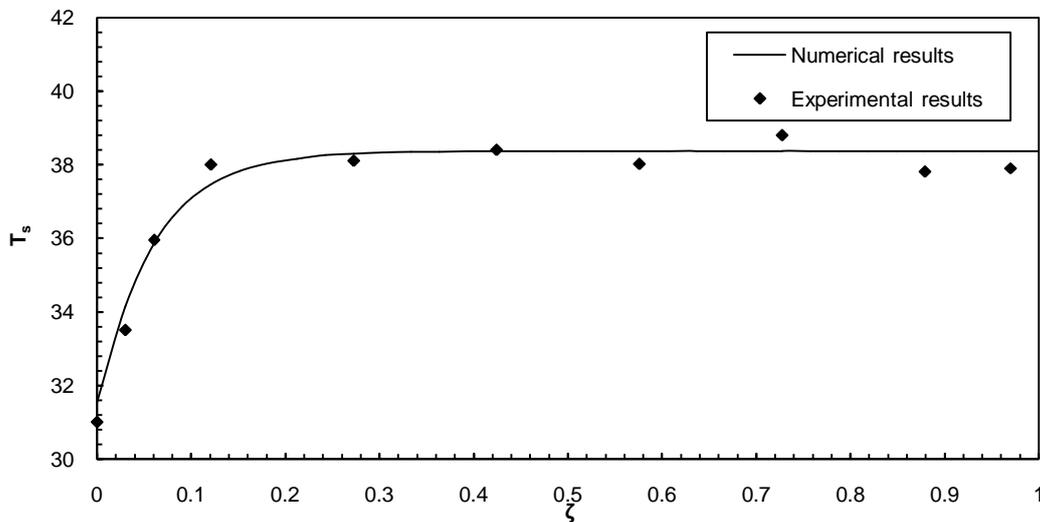


Fig. 10. Comparisons of the calculated and measured temperatures on the surface of the kharg oil pipeline in august 2008.

Heavy Metal Content and Physicochemical Properties of Municipal Solid Waste Dump Soils in Owerri Imo State

Eneje, Roseta. C.,¹ and Lemoha Kelechi, T²

¹Department of Soil Science And Agro-climatology, Michael Okpara University of Agriculture Umudike, Nigeria.

ABSTRACT: A study was carried out in Owerri municipal, the regional capital of Imo State to characterize the heavy metal content and physicochemical properties of waste dump sites for more than 15 years. Soil samples were collected from two different locations (Otamiri hilltop dumpsite and Otamiri Gully dumpsites) with control samples collected ten meters away from the dumpsites. The samples were collected at three depths (0-20cm, 20-40cm and 40-60cm) from each of the sites with three replicates making a total of 27 samples. Samples were taken for laboratory analyses for properties such as heavy metals (Ld, Fe, cu), chemical properties (organic carbon and Nitrogen pH CEC and EA) and physical properties (As, BD Ksat Particle size). Results show that the wastes dump sites showed variability in soil properties with depth the soils of the non-dump site at varying depth are classified as slightly acidic of high aggregate stability (76%). Comparison of the two sites indicated higher values in the soils properties for samples collected from gully dumpsite than those collected from hill top. Heavy metal content were generally higher at deeper depths (40-60cm) and the hill top waste dump site had higher values compared to the gully dump site only at shallow depth (0-40cm). Statistically, there was a significant ($p < 0.05$) difference in heavy metal content, exchangeable cations and soil pH as well as and BD and AS between two sites. Generally, the solid wastes increased the values for soil pH, CEC, heavy metals, aggregate stability, organic matter and total nitrogen when compared to adjacent uncontaminated soil. This is attributable to the decomposition and mineralization of the biodegradable solid wastes in the sites leading to release of minerals as well as basic cations into the soil which caused increases in soil physicochemical properties.

Keywords: Municipal solid waste, Owerri metropolis, heavy metal, soil pH, aggregate stability

I. INTRODUCTION

Solid wastes other than hazardous and radioactive materials are often referred to as municipal solid waste (MSW). According to Nyangabobo and Hamya, 1980, municipal solid waste are useless unwanted materials discharged as a result of human activity which commonly may be solids, semi solids or liquids in containers thrown out of houses, commercials or industrial premises. They are commonly called trash or garbage and include wastes such as durable goods, e.g. tires, furniture; non durable goods, e.g. newspapers, plastics wrap; and other wastes, e.g. yard waste, food etc. this category waste generally refers to common household wastes, as well as office and retail wastes but excludes industrial, hazardous and construction wastes. MSW varies in composition which may be influenced by many factors such as culture affluence, location and its management depends on the characteristic of the solid waste including the gross composition, moisture contents, average particle size, chemical composition and density. The United States environmental protection Agency defined solid waste as “any useless, unwanted or discarded material with insufficient liquid content to be free flowing”. The non-free flowing or sticky nature of the solid waste gives rise to the accumulation of solid wastes on some habitable parts of the earth surface, places with accumulated solid wastes are called refuse dumps but a designed place for dumping of refuse is known as dump site. Soils intensively affected by human activities might present special features such as mixed horizons, foreign materials and thin topsoil (Civeira and Lavado, 2006). Normally these soils are poor in organic matter and fertility with reductions in their most important physical properties, such as structural stability and water retention. Eventually, these characteristics have detrimental effects on the soil by either affecting plant growth or submitting the particular environment to erosion processes (Vetterlein and Hiittl, 1999; Scharenbroch *et al*, 2005). Several waste dumpsites are located at various parts of Owerri municipal and some of these sites are indiscriminately located at stream valleys; open fields water canals and in abandoned borrow pits. Studies by Ezeigbo (1989), Egboka and Umah (1985) show that there is an unconfined water waste aquifer underlying most of Owerri and environs, upon which all the divers depend for their various water needs.

Many management techniques of waste disposal abound but the most favorite techniques used in Owerri municipal, probably because of its low cost, is the disposal on land or in holes made in the earth. Modifications of these have popularized in Landfills, pit latrines and deep-well injections among others. One critical issue in all the management techniques is how to keep the harmful effects of waste away from man. A major concern in waste dumping either above or below the earth's surface is the safety of the ground water in the area. In areas of shallow water table and abundant rainfall the danger is greater since waste contaminated water reaches the aquifer easily. The resulting contamination can spread as far as the ground water can reach, sometimes through the entire aquifer. When an entire aquifer is polluted, the health of millions becomes endangered. Many designs have been employed in the construction of landfills with the main purpose being to fortify the walls, the bottoms of the pits and cover the top from contaminating the atmosphere. The bottom and the walls, however, are the contacts that conduct waste leachates into ground water (Allen, 2001). O'Swallwan (1995) observed that no matter the degree of structuring or engineering a landfill site receives, that only the natural local geology and hydrogeology of the site can guarantee its efficiency. Therefore the objective of his study is; to access the effect of soil waste dump on heavy metal content and compare it with those of adjacent land, and assess the effects of municipal solid waste on some selected physiochemical properties of the soil at varying depths.

II. MATERIALS AND METHODS SITE DESCRIPTION

This study was carried out in Owerri municipal, the regional capital of Imo State. It is located on latitude 5°N to 6°N and longitude 6° to 7° 34'E. It experiences two distinct climate seasons; namely dry seasons (October-March) and wet (April-September) seasons. A period of cold, dry, dusty winds known as "Harmattan" occurs from (December-February) annually. Owerri has a mean temperature range between 24° to 34°C with a relative humidity of 70% in dry months and 90% in wet months with a projected 2010 population of 610,211 people (NPC, 2007) unevenly distributed over a total land area of 50,885km², its residents are mainly traders, civil servants, artisans and blue collar workers in small-scale industries (Bakeries, food processing, medical laboratories, printing etc), a well developed network of major roads, access roads and streets also exist in the town.

The waste dump sites had been under used for more than 15 years, with a surface area of approximately six hectares. Nearly ten tons of wastes are dumped here each day, waste components include metals, (beverages, cans, ferrous materials etc.), used papers, rags, plastics and organic materials (food remnants, decaying leaves, fruits and vegetables etc) car batteries etc.

III. SAMPLE COLLECTION

Soil samples were collected from two different locations of the dumpsites namely Otamiri hilltop dumpsite and Otamiri Gully dumpsites both situated at Owerri-Aba road in Owerri municipality. Control sample were also collected at 10 meters away from the dumpsites. The samples were collected at intervals of 0-20, 20-40, 40-60cm from each of the sites with three replicates making a total of 27 samples. Samples were taken for laboratory analyses after picking out bulky materials and sieving with 5mm mesh.

IV. DETERMINATION OF HEAVY METAL CONTENT

This was determined using the digestive method of A.O.A.C (1975). 1g of dried and homogenized soil obtained from each samples were weighed into a 100ml beaker and 10ml of nitric acid was added and the mixture was reacted. The heating continued, followed by the addition of 10ml of HNO₃, 3ml of HClO₄ at intervals. HClO₄ was added and the solution filtered and diluted with water to 50ml mark. Standard solution of Lead (Pb), Zinc (Zn) and Iron (Fe) were prepared. The concentrations of the heavy metals were determined by using atomic absorption spectrophotometer. Other properties analysed include particle size analysis determined using the hydrometer method of Bouyoucous (1951), aggregate stability was determined using the mean weight diameter method as described by Kemper and Chepil, (1965) and soil pH was measured electrometrically with a glass electrode pH meter in KCL using a soil: liquid suspension ratio of 1:2:5 as modified by Jones (2001). Total nitrogen was determined by the method of Bremner and Malvaney (1992), while organic matter was determined using dichromate wet oxidation method (Walkley and Black, 1934) the organic carbon was calculated as: % organic carbon in soil = $\frac{(\text{MeK}_2\text{Cr}_2\text{O}_7 - \text{MeFeSO}_4) \times 0.003 \times 100 \times 1.33(F)}{\text{Gram of air dry soil}}$

Where; F = Correction factor

Me = Normality of solution x 1ml of solution used.

Exchangeable cations were extracted using ammonium acetate method, potassium, sodium were determined on a flame photometer while calcium and magnesium were determined by titrating with EDTA (Chapman, 1965).

V. STATISTICAL ANALYSIS

Soil data were subjected to analysis of variance using 2×3×3 factorial in RCBD and significant treatment means were separated by F_{LSD} 0.05.

VI. RESULTS AND DISCUSSIONS

Analysis of soils collected from site ten meters away from the wastes dump sites (Table1) showed variability in soil properties with depth, while soil pH, O.M and total nitrogen values decreased with increase in depth, CEC, and some heavy metals values were inconsistent across the sampling depth. However, the soils of the non dump site at varying depth are classified as slightly acidic of high aggregate stability (76%).

Table 2 and 3 shows the physicochemical properties of soil collected from the solid waste dump sites at varying depths. These results indicate that solid wastes had a significant effect on all the soil properties, these effects varied significantly (P < 0.05) with depth. Generally, the solid wastes increased the values for soil pH, CEC, heavy metals, aggregate stability, organic matter and total nitrogen when compared to adjacent uncontaminated soil. This is attributable to the decomposition and mineralization of the biodegradable solid wastes in the sites leading to release of minerals as well as basic cations into the soil which caused increases in soil physicochemical properties.

Comparison of the two sites indicated higher values in the soils properties for samples collected from gully dumpsite than those collected from hill top. Heavy metal content of the sampled soils (Figs.1, 2, and 3) indicated that the metals were generally higher at deeper depths (40-60cm) and the hill top waste dump site had higher values compared to the gully dump site only at shallow depth (0-40cm). Statistically, there was a significant difference between two sites and increases in the values of soil properties analyzed for the valley dumpsite compared to that of hill top could be explained from differences in the topography of sites such that most materials are moved from the hill top to the valley by erosion.

VII. SUMMARY AND CONCLUSION

The concentration of heavy metals, total nitrogen, organic matter, soil pH, and cation exchange capacity were observed to be higher in soils at dumpsite compared to those obtained in adjacent soils 10 meters away. This implies that the

municipal solid waste dumpsite have a significant impact on the environment. From this study, the municipal solid waste impacted or decreased the soil pH, OM and CEC down the depth, while heavy metals deposit increased down the depth. Also, the soil properties studied had higher values in the soils collected from valley dump site than those collected from hill top.

Table1: Properties of Soil Collected From An Uncontaminated Site At Different Depths

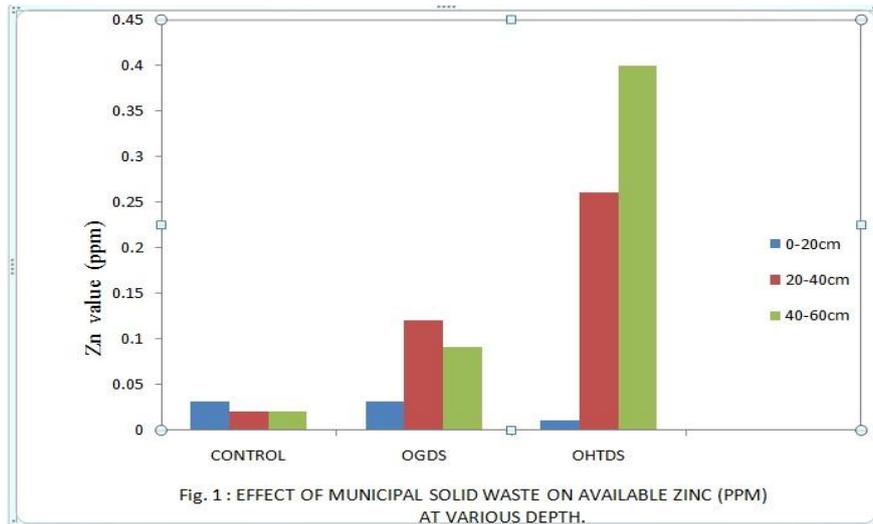
Soil Parameters	Soil Depths		
	0-20	20-40	40-60
Particle Size Distribution			
% Sand	80.0	71.40	68.00
% Silt	2.90	3.60	4.20
% Clay	17.10	25.00	27.80
pH (H ₂ O)	6.34	6.26	6.14
Organic Matter Content (%)	1.24	0.90	0.78
Total Nitrogen (%)	0.098	0.058	0.47
Cation exchange Capacity (Cmol.kg ⁻¹)	3.86	3.72	3.96
Heavy Metals			
Zinc	0.03	0.02	0.02
Iron	0.36	0.20	1.00
Lead	0.04	0.07	0.02
Aggregate stability (%MWD)	80.40	53.10	73.70

Table 2: Effect of Solid Waste on Some Properties of Soil Collected at the Valley Dumpsite at varying depths

Soil Parameters	Soil Depths			L.S.D(0.05)
	0-20	20-40	40-60	
Particle Size Distribution				
% Sand	75.40	72.50	70.02	-
% Silt	3.30	5.60	6.50	-
% Clay	21.30	21.90	23.48	-
pH (H ₂ O)	7.04	6.58	6.39	0.461
Organic Matter Content (%)	2.26	1.47	0.98	0.180
Total Nitrogen (%)	0.23	0.17	0.09	0.039
Cation exchange Capacity (Cmol.kg ⁻¹)	5.05	4.64	4.88	0.461
Heavy Metals				
Zinc	0.03	0.12	0.09	0.038
Iron	0.30	1.80	0.70	0.278
Lead	0.22	0.23	1.12	0.091
Aggregate stability (%MWD)	85.5	7.3	82.2	7.968

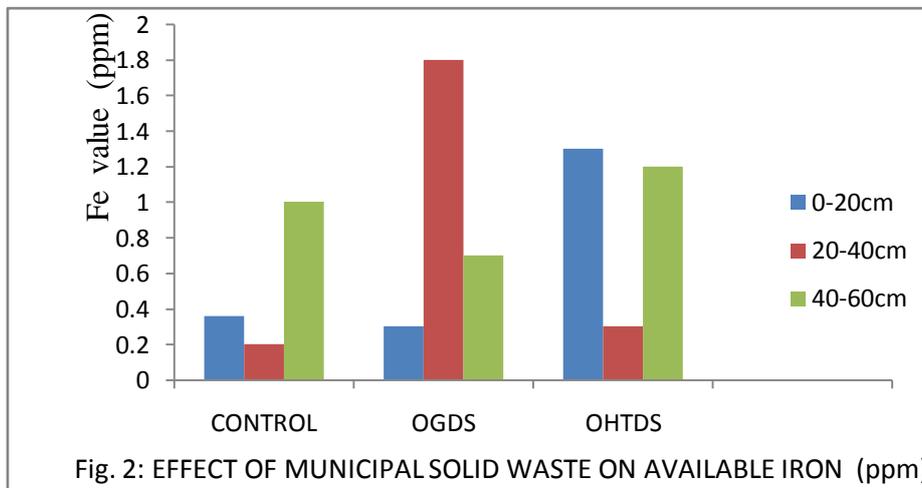
Table 3: Effect of Municipal Solid Waste on some Properties of Soil collected at the Hilltop Dumpsite.

Soil Parameters	Soil Depths			L.S.D(0.05)
	0-20	20-40	40-60	
Particle Size Distribution				
% Sand	74.8	69.4	64.2	-
% Silt	2.6	2.8	4.6	-
% Clay	22.6	27.8	31.2	-
pH (H ₂ O)	7.37	7.22	6.80	0.46
Organic Matter Content (%)	1.84	1.20	1.08	0.18
Total Nitrogen (%)	0.19	0.09	0.06	0.04
Cation exchange Capacity (Cmol.kg ⁻¹)	4.12	3.88	3.79	0.46
Heavy Metals				
Zinc	0.01	0.26	0.40	0.04
Iron	1.30	0.30	1.20	0.28
Lead	0.46	0.32	0.93	0.10
Aggregate stability (%MWD)	40	25.2	35.9	7.97



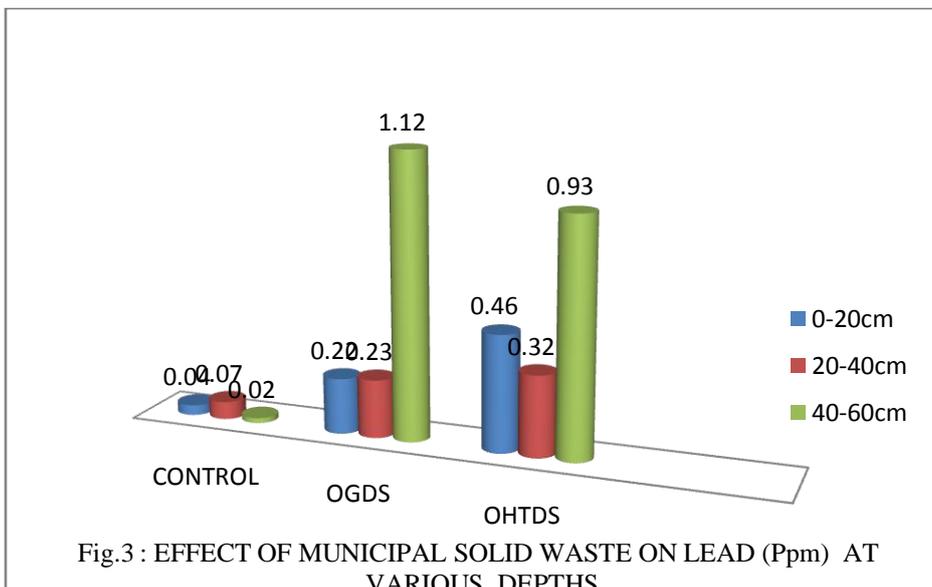
OGDS – OTAMIRI GULLY DUMP SITE

OHTDS – OTAMIRI HILL TOP DUMP SITE



OGDS – OTAMIRI GULLY DUMP SITE

OHTDS – OTAMIRI HILL TOP DUMP SITE



OGDS – OTAMIRI GULLY DUMP SITE

OHTDS – OTAMIRI HILL TOP DUMP SITE

REFERENCES

1. A.O.A.C (1975). Official Methods of Analysis, association of official analytical Chemists, 12th Edition. William Horwitz, Pp.9-12.
2. Allen R. J., (2001). Household Hazardous waste in Municipal Landfills: Contaminant in Leachate. Sci. Total Environ. 337:119-137.
3. Bouyoucos, G.J (1951). A Recalibration of the Hydrometer for Making Mechanical Analysis of Soils. *Agronomy journal* 43:434-438.
4. Bremner, J.M and Mulvaney (1982). Determination of Nitrogen in Soil by the Kjeldahl Method. *Journal of Agricultural Science*. 55:11-33.
5. Chapman H.D., (1965). Total Exchangeable Base in; Black, C.A (ed). Methods of Analysis Part 11.am, Soc, Agron 9 Madison Pp.902-904.
6. Civeira G., and R.S Lavado., (2006). Nitrate Losses, Nutrients and Heavy Metals Accumulation from Substrates Assembled for Urban Soils Reconstruction. *Journal of Environmental Management* 88:1619-1623.
7. Egboka, B.C.E., Uma, K.O., (1985). New Statistical Grain-size Methods for Evaluating the Hydraulic Conductivity of Sandy Aquifers *J.Hydrol* 108:143- 366.
8. Ezeigbo, H.I., (1989). An Evaluation of the Ground Water resources of Ogbunike Area and Environs Anambra State. S. E. Nigeria. *Water Res. J.16(1&2)*.
9. Kemper, W.D., Chapil, W.S., (1965). Size distribution of Aggregate. In: Black, C.A., (ed). *Methods of Soil Analysis. Part I Am.Soc. Agron.9, Madison, Wisconsin Pp. 499- 510*.
10. Nyangababo, J.T., Hamya J.W., (1980). The deposition of Lead, Zinc, Cadmium and Copper from Motor Traffic on Brandaria Enimi and Soil along a Major Bombo Road. *Int. J. Environ. Stud. Legol* 1(32):117.
11. O'swallan, S.W., (1995). Control of Municipal Solid Wastes. Oxford University press Oxford 6th eds. P.642.
12. Scharenbroch, B.C., J.E, Lloyd, and J.L, Johnson., (2005). Distinguishing Urban Soils with Physical, Chemical and Biological Properties. *Pedobiologia* 49:283-296.
13. Vetterlein, J., and Hiittl, U., (1999). Can Applied organic Matter Fulfill Similar Functions as Soil Organic Matter, Risk Benefit Analysis for Organic matter Application as a Potential Strategy for rehabilitation of Disturbed Ecosystems. *Plant and Soil* 213:1-10.
14. Walkley, A., and Black, I.A., (1934). An Examination of the Depth-Jare Method for Determining Soil Organic Matter and Proposed Modification Method. *Soil Science* 37:29-38.

Industrial Process Heat Applications of Solar Energy

A.G. Bhave

Department of Mechanical Engineering, K.J. Somaiya College of Engineering, Vidyavihar (E), Mumbai- 400077, India

ABSTRACT: Concentrating solar collectors of low concentration ratio can supply hot water or process heat at intermediate temperatures, i.e. 80 to 180 °C, where flat plate collectors are not effective, and for which there are a number of industrial, domestic and rural applications. This paper studies the potential applications of this technology and presents case studies to prove its feasibility.

Keywords: Concentrating solar collector; cooling; desalination; industrial process heat; water pumping

I. INTRODUCTION

Large amounts of energy are spent for industrial heat generation in many countries. India uses 100 million tons of oil annually, of which 40 % is consumed in industries. 60 – 70 % of this use is in the form of thermal energy, out of which 70 % is used for applications below 250 °C. 30 % of this requirement can be met through solar concentrating collectors, leading to savings of about 4.5 million tons of furnace oil, LDO or diesel [1]. India has a large number of small, decentralized industrial units, where this technology can be used, leading at the same time to a reduction in oil imports. The development of solar thermal systems for industrial heat and for other applications like solar cooling is a thrust area of the Ministry of New and Renewable Energy, Government of India. This paper studies the potential application areas for concentrating solar collectors and presents a review of actual operating systems or research work proving potential application areas.

II. APPLICATIONS OF SOLAR ENERGY IN MEDIUM TEMPERATURE RANGE

There are several potential fields of application of solar thermal energy in the temperature range of 60°C to 180 °C, like heat production for industrial processes in the food, textile, wine and chemical industries, solar cooling and air conditioning, solar drying and seawater desalination, thermal detoxification of drinking water, and small power generation and water pumping through medium temperature Rankine cycle systems.

An overview of the potential of solar industrial process heat in the food industry is presented below [2]. The applications require heat in the temperature range 60 – 180 °C:

<u>Industry</u>	<u>Process</u>	<u>Temp. Range, °C</u>
Dairy	Pasteurisation	60 - 85
	Sterilisation	130 - 150
	Drying of milk powder	60 - 180
Food Preservation	Sterilisation of vegetables, fish, meat, baby food	Hot water at 130- 150 °C or steam
	Cooling by solar energy	Hot water or steam
	Meat	Washing
	Sterilisation	Upto 100

Worldwide energy consumption for cooling and air-conditioning is rising rapidly and the market potential for solar thermal cooling is very large. Solar energy has the advantage that cooling is generally required when solar radiation is available. This is the main reason for sustained research into solar cooling devices for at least three decades. These studies include solar energy technologies operating with absorption, adsorption, and desiccant cycles to produce cooling and refrigeration using medium to high-temperature solar technologies (from 80 to 250 °C)

However, the market is still very much at an initial stage, with only around 500 solar cooling systems installed globally, and has been largely dominated by Europe so far [3].

Small scale water pumping systems are widely used in India for pumping drinking and irrigation water from wells, and for lifting drinking water to roof-based tanks in urban housing societies. This is accomplished by electrical or diesel-engine driven pumps, or in case of renewable energy systems, by solar photovoltaic or biomass gasifier- IC engine pumpsets. The area of solar thermal water pumping systems has been the subject of development work by researchers around the world [4], though such systems are not widely used in India. Stationary concentrators like the CPC give the option of operating a Rankine cycle system at temperatures above that permitted by flat-plate solar collectors.

III. CONCENTRATING SOLAR COLLECTORS

Solar concentrating collectors are used to increase the incident flux of sunlight on a receiver which has a smaller area than the aperture of the collector. Due to the smaller surface area of the receiver, it has a lower heat loss, and therefore reaches a higher temperature than non-concentrating collectors. They use reflective surfaces to reflect and concentrate sunlight onto a

small area, where it is absorbed and used as heat or, in the case of solar photovoltaic devices, it is converted into electricity. Refractive media like Fresnel lenses are also used for the concentration of solar thermal energy, and concentrators using Fresnel lenses in conjunction with photovoltaic cells are also under development.

Concentrating collectors can be used for high-temperature applications such as steam production for power generation, for industrial process heat and other applications.

One of the most well known concentrators is the parabolic trough collector or cylindrical parabolic concentrator (Fig. 1). It concentrates beam radiation of the sun onto a tubular receiver-absorber at the focus of the parabolic mirror. The tube carries the heat transfer fluid. This system therefore needs a clear sky as well as tracking of the movement of the sun during the daytime.



Fig. 1 Cylindrical parabolic concentrating solar collector

Compound parabolic concentrating solar collectors [5] (CPC) are non-imaging concentrators, which consist of two reflectors which are segments of two parabolas, which focus the sunlight entering its aperture onto a smaller, flat or round absorber surface, where it heats up the working fluid (Fig.2).

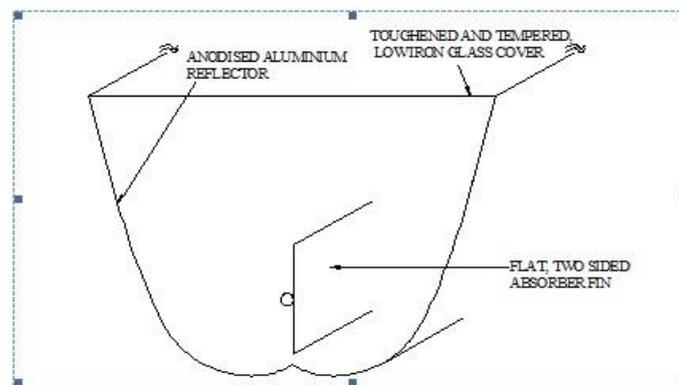


Fig. 2 CPC solar collector with vertical absorber

CPCs of low concentration ratio can supply hot water or process heat at intermediate temperatures, i.e. 80 to 180 °C, where flat plate collectors are not effective, and for which there are a number of applications. As they collect both beam and diffuse radiation, they can also be used under conditions of low insolation. They can be made at a relatively low cost as they are non-tracking devices when mounted east-west, and can be used with seasonal adjustment of the tilt angle.

Parabolic dish concentrators focus the beam radiation of the sun onto a point, and therefore require continuous, 2-axes tracking of the sun to keep its aperture plane perpendicular to the beam of the sun. They can be used with high concentration ratios to achieve temperatures of more than 400 °C at the focus.

Concentrating collectors are generally not used in domestic applications. An exception is the solar cooker based on a parabolic dish concentrator, used for community cooking.

IV. CASE STUDIES ON APPLICATION AND RESEARCH ON SOLAR CONCENTRATING COLLECTORS

Cases of research on and thermal applications of concentrating solar collectors are covered in the literature survey below.

4.1 Industrial process heat:

Solar concentrating collectors can be retrofitted to the existing boiler or heater system in the industry, substituting the use of furnace oil or electricity used for the heating application partially or fully.

One example of such an application is the 160 m² paraboloid concentrating dish based system installed at Mahanand Dairy, Latur by Clique Consultants Pvt. Ltd., Mumbai [6]. This system generates hot water at 180 °C and 18 bar pressure, to be used for milk pasteurization. It saves 70 – 90 litres of furnace oil on a clear sunny day.



Fig. 3: ARUN solar concentrator system, Clique Consultants

4.2 Desalination:

A study done in Spain [7] involves an experimental hybrid solar-gas fired seawater desalination plant based on a multi-effect distillation process (72 m³/day distillate), with thermal input in the temperature range 60 – 70 °C. Part of its energy comes from a static solar field of stationary CPC collectors (500 m² surface area). Water is the heat transfer fluid for the collector field. The hot water is stored in storage tanks. The system can operate in one of three modes:

- Solar only mode: energy to the first distillation effect comes exclusively from the solar collector field
- Fossil only mode: the gas boiler operated double effect absorption heat pump supplies all the heat required by the distillation plant
- Hybrid mode: the energy comes from both the heat pump and the solar field

To optimize the plant operation, the authors have done a modelling study, validated by experimental data.

4.3 Cooling:

A study done for the hot climate of Nicosia, Cyprus [8], presents the optimisation of the various components of a lithium bromide-water absorption solar cooling system with a generator temperature of 75 °C, such as the type, slope and area of solar collector and storage tank size, for a house. The collector types considered by the authors are the flat plate, compound parabolic and evacuated tube collectors. The optimisation considers the amount of useful energy collected against the life cycle cost of the solar system, and is carried out using TRNSYS software. The solar system can be used during summer to provide part of the heat required by the absorption cooling unit and during winter to provide part of the heating load. The results indicate that due to the high cost of fuel a large part of the building load can be covered by solar energy. They found that the optimum system is one based on a CPC collector array.

4.4 Commercial CPC solar collectors:

Solargenix Energy [9], USA, offers their Winston Series CPC based solar collectors for use for hot water production, space heating and solar cooling applications requiring heat at upto 98 °C. It is a single glazed, non-evacuated collector with selectively coated tubular absorbers.

V. CONCLUSIONS

The literature survey presented above leads to the following conclusions:

- i) There is good potential for replacement of petroleum derived oils and conventional electric power presently being used for industrial process heat applications, power generation and applications like cooling and water desalination, with solar energy by the use of concentrating solar collectors.
- ii) A large number of these applications are likely to be at a small scale, for process heat in small and medium industries, for domestic cooling and drinking water production and for small scale agricultural water pumping.
- iii) The studies described show that technically feasible solutions, which are economical under certain conditions, are available.

REFERENCES

- [1] Shireesh Kedare, Solar thermal systems for industrial process applications, *RENET Workshop. IIT Bombay, 21 – 22 September, 2006.*
- [2] Soteris Kalogirou, The potential of solar energy in food-industry process heat applications, *25th National Renewable Energy Convention. Warangal, India, 2001.*
- [3] David Appleyard, Chilling out in the sun: solar cooling, *Renewable Energy World International Magazine. 13 Issue 3 (2010).*
- [4] A.G. Bhave, Potential for solar water- pumping systems in India, *Applied Energy. 48 (1994), 197 – 200.*
- [5] Ari Rabl, Optical and thermal properties of compound parabolic concentrators, *Solar Energy. 18 (1976) 497 - 511.*
- [6] Website of Clique Developments Ltd, Mumbai – www.clique.in
- [7] L. Roca, L.J. Yebra, M. Berenguel, Diego C. Alarcon-Padilla, Modeling of a solar sea water desalination plant for automatic operation purposes, *Journal of Solar Energy Engineering. 130 (November 2008) 041009-1 to 041009-8.*
- [8] Georgios A. Florides and Soteris A. Kalogirou, Optimisation and cost analysis of a lithium bromide absorption solar cooling system, *Proceedings of Clima 2007 WellBeing Indoors. Helsinki, Finland, 10 – 14 June 2007.*
- [9] Website of Solargenix Energy, USA, www.solargenix.com

Elimination of Heart Signal Baseline Wandering by Neural Network

Alireza Akoushideh¹, Ali Zakerolhosseini², Babak Mazloom-nezhad Maybodi³,
Mohammadtaghy Mirzaee⁴

²(Electrical and Computer Department, University of Shahid Beheshti G.C, Tehran, Iran

¹(Shahid-Chamran Junior College, Technical and Vocational University, Rasht, Iran

³(Mirza-Koochak Department, Technical and Vocational University, Somesara, Iran

ABSTRACT : In this paper, we propose artificial neural network approach for elimination of Electro Cardio Gram (ECG) baseline wandering. This technique uses self-organization feature map (SOM) technique without any disorder in important parameters such as ST-segment. Proposed technique not only eliminates base line divergence better than classic methods but also have lower disorder in ST-segment.

Keywords: ECG, SOM, neural network, baseline noise, ST-Segment.

I. INTRODUCTION

In Cardiac Care Unit (CCU) nurse should pay attention to vital signals (heart, brain, etc.) of several patient continuously. Certainly supervising of them and quick decision-making will be difficult. Therefore it is necessary that an intelligent system contribute to human in decision-making. Noises in heart signals could be recognized by experienced doctors and they could analyze ECG signal without any problem. However, noise is so much important in correct diagnosis of disease with assistant machine. To recognize the heart function, it is essential that we know facts of heart physical activity which is possible by receiving of heart electrical signals which come to body surface [1]. Our object in this research is elimination of ECG baseline to better analyze of this critical wave.

This pre-processing is necessary for other essential processing to work correctly. These algorithm are such as ECG signal compression[2, 3], QRS complex detection[4, 5], heart rate Variability[6], ST segment analysis[7], arrhythmia diagnosing[8] and even arrhythmic elimination (by pace makers), Morphological Characteristics of P-Waves[9], Tele monitoring[10, 11], could be available by time analyses method, neural network method[12, 13], syntactic method[14], and Hidden Markov model[15]. Therefore, noises should be cleared from signal.

In this article, our contribution is decreasing of ECG baseline and comparison to classical techniques. This paper is organized as follows. In section II, two important types of noise in biomedical measuring and appraise current methods of baseline elimination are introduced. General reviewing of self-organization feature map is in section III. In section IV proposed algorithm are described. Experimental results of proposed method are shown in section V and conclusion is in section VI.

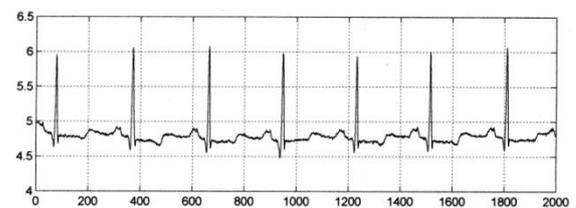
II. BASE LINE WANDERING NOISE AND ELIMINATION METHODS

A. Different types of noises

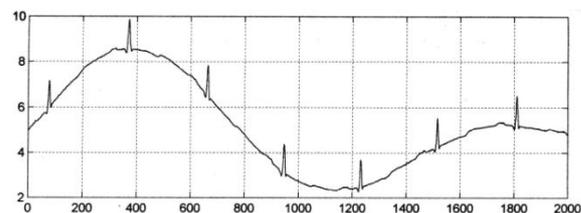
Noises in heart signal are divided into two groups in frequency.

- Disorders which have higher frequency compare with ECG signal. Disorder source alters signal by undesirable pulses which are called impulsive.
- Disorders which have lower frequency and could make slow changes in signal. This group of noises are called baseline wandering.

To eliminate of these kinds of noises, there are different methods such as time-based method [16], frequency-based method, neural network methods and so on. Fig. 1 shows a Normal ECG and a base line noisy signal.



(a) Normal ECG



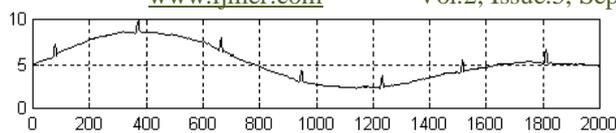
(b) ECG wave with base wandering noise

Fig. 1. Normal and base line wandering noise

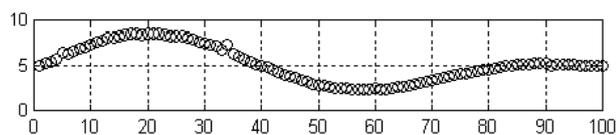
A. Base line elimination methods

Using of a low pass filter and a filter which could eliminate AC power frequency, remove of high frequency noise is not problematic. But for elimination of base line wandering, which has alternative and partial low frequency and unpredictable extent, find of an effective solution is necessary. Available base line wandering elimination methods include:

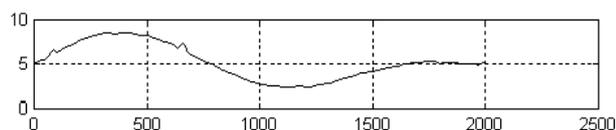
- Method I: getting of low rate samples from the wave and pass a curve (e.g. spline curve) through these points. The curve makes base line of wave, which by subtraction of ECG wave noise free ECG wave will be made. Figure 2(a) depicts an ECG wave with 2000 points. Low rate sampling, getting 100 points (5%), is shown in Fig. 2(b) and Fig 2(c) is predicting of base line. Figure 2(d) shows this method is not efficient. Because, some sample point are get from local picks of wave and it can not makes a noise free wave.



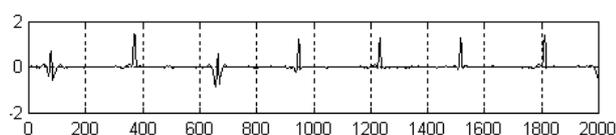
(a) Normal ECG signal with baseline noise



(b) Low sampling (5%)



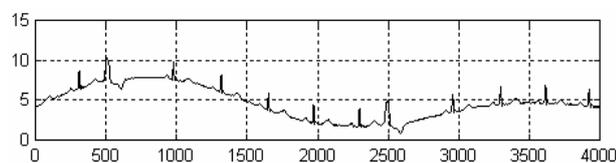
(c) Baseline estimation curve



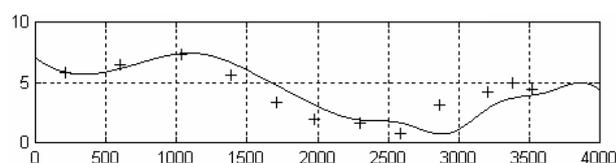
(d) Removing of baseline

Fig. 2. Baseline noise damping, Method I

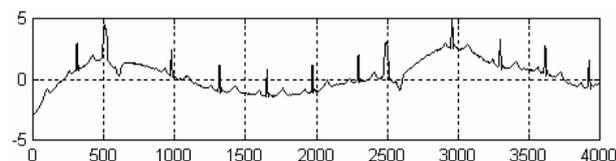
- Method II: This technique find beginning point of QRS complexes and then passes a curve (e.g. spline curve) through the obtained points. Other processing is similar to method I. Fig. 3(a) to 3(c) depict process of this technique. Figures show that this technique for abnormal wave, such as PVC wave, is not efficient.



(a) Arrhythmic ECG signal with baseline noise



(b) Detection of ECG beat's position and plotting of baseline estimation curve



(c) Removing of baseline

Fig. 3. Base line noise damping, Method II

III. ARTIFICIAL NEURAL NETWORK METHODS [17]

In recent years, many kinds of information processing have not any classic solution and could not be solved easily. One of these modern and experience based methods is Artificial Neural Networks (ANN) which transfer hide knowledge and rule of data to network structure by processing of experimental data. These

intelligent systems learn general rules based on calculation of numeral data or examples and after training they could decide to solve a new problem and also they have very strong pattern recognition.

Competitive neural network as dependent detector of signal density

With random selection of inputs and apply them in network, after every repetition, winner cell weight vector moves towards input vector. Finally, each weight vector directs towards cluster point of some input vectors. In this competitive method, reference pattern is not available and weight vector of a cell is changed, in other words a row of weight matrix which is nearest to input vector tends to input vector. Therefore, each cell is represents a group of applied inputs.

Self-Organization Feature Map (SOM)

A. Introduction

A self-organizing map (SOM) or self-organizing feature map (SOFM) is a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional (typically two-dimensional), discretized representation of the input space of the training samples, called a map. Self-organizing maps are different from other artificial neural networks in the sense that they use a neighbourhood function to preserve the topological properties of the input space.

B. Learning algorithm

The training utilizes competitive learning. When a training example is fed to the network, its Euclidean distance to all weight vectors is computed. The neuron with weight vector most similar to the input is called the best matching unit (BMU). The weights of the BMU and neurons close to it in the SOM lattice are adjusted towards the input vector. The magnitude of the change decreases with time and with distance from the BMU.

1. Randomize the map's nodes' weight vectors
2. Grab an input vector
3. Traverse each node in the map
 - a. Use Euclidean distance formula to find similarity between the input vector and the map's node's weight vector
 - b. Track the node that produces the smallest distance (this node is the best matching unit, BMU)
4. Update the nodes around BMU by pulling them closer to the input vector according to following by (1):

$$W_v(t+1) = W_v(t) + \Theta(t)\alpha(t)(D(t) - W_v(t)) \quad (1)$$

Where, t denotes current iteration, W_v is the current weight vector, D is the target input, $\Theta(t)$ is restraint due to distance from BMU, usually called the neighbourhood function, and α is learning restraint due to time.

5. Increase t and repeat from 2 while $t < \lambda$

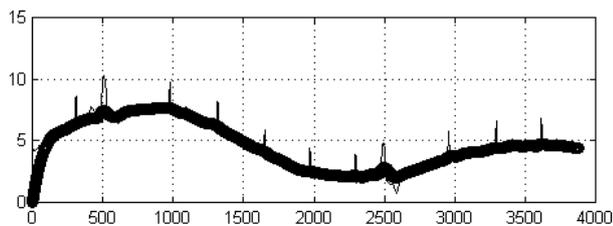
Where, λ is the limit on time iteration.

IV. DESCRIPTION OF PROPOSED METHOD

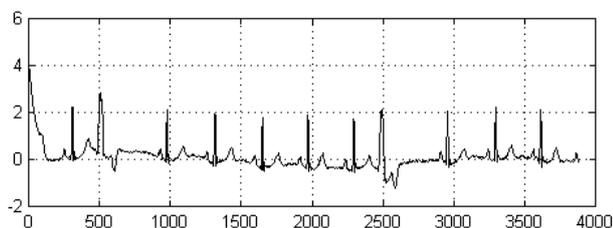
The main idea of this method is based on density of baseline, signal pause parts, which has more density than other part of ECG wave such as P, QRS complex and T waves. This means that quick and short time changes (such as QRS complex) could not allocate high quality of signal density. Therefore, after training by SOM algorithm, each

neuron represents several near points of signal curve. Clearly, Fig. 4 depicts that, in a competitive system distribution of cells around the baseline is more than other part of wave.

This cells estimate baseline curve of ECG wave and by subtraction of noisy image from predicted baseline, elimination of baseline wandering without any damaging on other part of ECG wave, specially ST-segment, are attained. Figure 4 shows proposed noise elimination that its quality can be compared with other methods. For implementation of SOM, ratio of neurons number to wave vector dimension chose 1 to 40.



(a) Tracing of base line by neurons in SOM algorithm

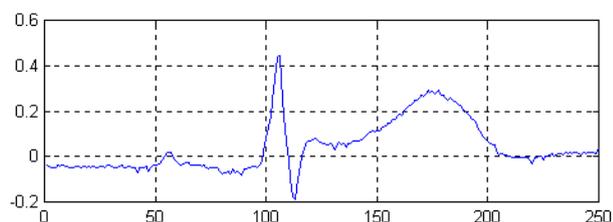


(b) Tracing of baseline by trained neurons in SOM and finding of base line curve

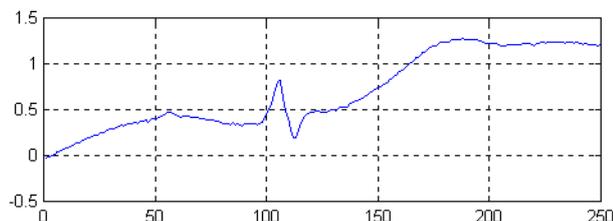
Fig. 4. Recovery of baseline by SOM method; proposed technique

V. EXPERIMENTAL RESULTS

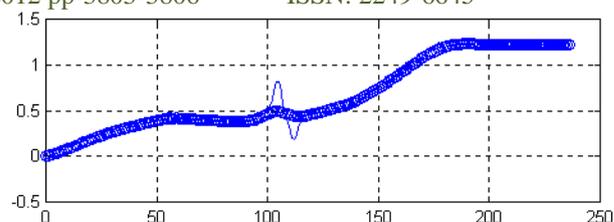
In order to show this method does not make any damaging on ST-segment, this technique was applied to one period of baseline wandered heart signal. Ratio of neurons was chose 1 to 10. Fig. 5(a) and 5(b) show normal and baseline noisy wave. Figure 5(c) depicts tracing of ECG wave baseline by neurons in SOM method. Finally, Fig. 5(d) depicts elimination of baseline wandering with least effect on ST segment according to other methods.



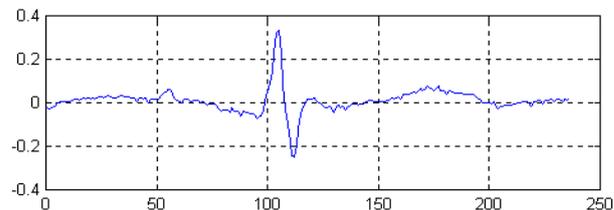
(a) One period normal ECG wave



(b) Base line noisy signal



(c) Tracing of baseline by trained neurons in SOM and finding of baseline curve



(d) Base line noise elimination with no damaged ST-segment

Fig. 5 Noise cancelation process in proposed method

VI. CONCLUSION

Method that we proposed in this paper is based on self-organization feature map (SOM) technique in neural network. Comparison with other methods it makes better removing of baseline wandering with least effect on ST segment. Furthermore, ratio 1:40 for neuron number to signal vector is enough for removing of baseline wandering.

VII. ACKNOWLEDGMENT

The authors wish to thank Dr. M.B. Menhaj and Dr. M. A. H. Fassihi for their assistance, useful discussions and support.

REFERENCES

- [1] B. Furht and A. Perez, "An adaptive real-time ECG compression algorithm with variable threshold," *Biomedical Engineering, IEEE Transactions on*, vol. 35, pp. 489-494, 1988.
- [2] K. Hyejung, R. F. Yazicioglu, P. Merken, C. Van Hoof, and Y. Hoi-Jun, "ECG Signal Compression and Classification Algorithm With Quad Level Vector for ECG Holter System," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, pp. 93-100, 2010.
- [3] B. R. S. Reddy and I. S. N. Murthy, "ECG Data Compression Using Fourier Descriptors," *Biomedical Engineering, IEEE Transactions on*, vol. BME-33, pp. 428-434, 1986.
- [4] N. M. Arzeno, D. Zhi-De, and P. Chi-Sang, "Analysis of First-Derivative Based QRS Detection Algorithms," *Biomedical Engineering, IEEE Transactions on*, vol. 55, pp. 478-484, 2008.
- [5] Y. Suzuki, "Self-organizing QRS-wave recognition in ECG using neural networks," *Neural Networks, IEEE Transactions on*, vol. 6, pp. 1469-1477, 1995.
- [6] S. Ahmad, M. Bolic, H. Dajani, V. Groza, I. Batkin, and S. Rajan, "Measurement of Heart Rate Variability Using an Oscillometric Blood Pressure Monitor," *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, pp. 2575-2590, 2010.
- [7] E. Skordalakis, "Recognition of the Shape of the ST Segment in ECG Waveforms," *Biomedical*

Engineering, IEEE Transactions on, vol. BME-33, pp. 972-974, 1986.

- [8] Zeybekog, x030C, S. lu, x, and M. zkan, "Classification of ECG Arrhythmia beats with Artificial Neural Networks," in Biomedical Engineering Meeting (BIYOMUT), 2010 15th National, 2010, pp. 1-4.
- [9] D. C. S. Teiichi Yamane, Jing-Tian Peng, Pierre Jaïs, Mélèze Hocini, Isabel Deisenhofer, Kee-Joon Choi, Laurent Macle, Jacques Clémenty, Michel Haïssaguerre "Morphological characteristics of Pwaves during selectivepulmonaryveinpacing," Journal of the American College of Cardiology, vol. 38, p. 6, 1 November 2001.
- [10] A. Alesanco, Garci, x, and J. a, "Clinical Assessment of Wireless ECG Transmission in Real-Time Cardiac Telemonitoring," Information Technology in Biomedicine, IEEE Transactions on, vol. 14, pp. 1144-1152, 2010.
- [11] S. P. Hariton N Costin, Cristian Rotariu, Bogdan Dionisie, Marinela C Cimpoesu, "A Multimedia Telemonitoring Network for Healthcare," Transactions on engineering, computing and technology, vol. 17, p. 6, 17 December 2006.
- [12] H. Atoui, J. Fayn, and P. Rubel, "A Novel Neural-Network Model for Deriving Standard 12-Lead ECGs From Serial Three-Lead ECGs: Application to Self-Care," Information Technology in Biomedicine, IEEE Transactions on, vol. 14, pp. 883-890, 2010.
- [13] S. Barro, M. Fernandez-Delgado, J. A. Vila-Sobrino, C. V. Regueiro, and E. Sanchez, "Classifying multichannel ECG patterns with an adaptive neural network," Engineering in Medicine and Biology Magazine, IEEE, vol. 17, pp. 45-55, 1998.
- [14] S. K. B. a. N. R. Manna, "Diagnosis of ECG diagnosable diseases- Syntatic Pattern Recognition Approach," in International Journal of Electronics vol. 68, ed, 2007, pp. 915-924.
- [15] F. Sandberg, M. Stridh, and L. Sornmo, "Frequency Tracking of Atrial Fibrillation Using Hidden Markov Models," Biomedical Engineering, IEEE Transactions on, vol. 55, pp. 502-511, 2008.
- [16] G. Jianbo, H. Sultan, H. Jing, and T. Wen-Wen, "Denoising Nonlinear Time Series by Adaptive Filtering and Wavelet Shrinkage: A Comparison," Signal Processing Letters, IEEE, vol. 17, pp. 237-240, 2010.
- [17] M. B. Menhaj, Fundamentals of Neural Networks: Amirkabir University of Technology, 2009.

An Efficient Approach for Packet Loss Measurement

Kanchana P¹, G. Nageswara Rao²

¹(M.Tech, Department of IT, Aditya Institute of Technology and Management, Andhra Pradesh, India)

²(professor & HOD, Department of IT, Aditya Institute of Technology and Management, Andhra Pradesh, India)

ABSTRACT: Various tools are available in market to measure packet loss in the end-to-end networks. But all the existing tools are not able to calculate packet loss in the network accurately. So measuring packet loss efficiently is still an open problem in the networks due to relatively occurrence and short duration of packet loss. The aim of our research is to understand how to efficiently calculate packet loss in the network. In this paper we presented an algorithm for measuring packet loss in the network. To evaluate the performance of the proposed algorithm, we simulate the algorithm in the NS2 simulator. Our experimental results shows the trade-offs between impact on the network and measurement accuracy. We show that measuring packet loss is more efficient in our proposed algorithm than traditional loss measurement tools.

Keywords: packet loss, NS2, Bernoulli loss model, BADABING

I. INTRODUCTION

Users can monitor network nodes for packet loss on routers using SNMP. Various tools are available in market to measure packet loss in the end-to-end networks. But all the existing tools are not able to calculate packet loss in the network accurately. PING is the most commonly used tool for measuring packet loss in the end-to-end paths. PING tool send ICMP echo packets to the destination at fixed intervals.

Sender assumes an occurrence of packet loss, if the acknowledgement from the destination is not received within a specified time period [1], [2]. Synchronization between sender and receiver is very important for measuring packet loss in the network. Strict synchronization of two entities connected by a varying delay link, can prove to be impossible without access to an external universal time reference as provided by a GPS (Global Positioning System) time reference. Even if GPS acquisition cards are now more frequently used enabling feasible delays with a resolution around 1 μ sec, It worths to try to extract as much as information from the loss process which much more simple to measure. A measurement approach is problematic because of the discrete sampling nature of the probe process. Thus, the accuracy of the resulting measurements depends both on the characteristics and interpretation of the sampling process as well as the characteristics of the underlying loss process.

Joel Sommers [3] proposed an approach in the network, tells us that Poisson modulated probes will provide unbiased time average measurements of a router queue's state. But this method needs higher moments of measurement to determine valid results. A closely related issue is the fact that loss is typically a rare event in the Internet [4]. This reality implies either that measurements must be taken over a long time period, or that average rates of Poisson-modulated probes may have to be quite high in

order to report accurate estimates in a timely fashion.

However, increasing the mean probe rate may lead to the situation that the probes themselves skew the results.

Thus, there are trade-offs in packet loss measurements between probe rate, assurance accuracy, impact on the path and timeliness of results [5], [6]. Measuring and analyzing network traffic dynamics between end hosts has provided the foundation for the development of many different network protocols and systems. Of particular importance is understanding packet loss behavior since loss can have a significant impact on the performance of both TCP- and UDP-based applications. Despite efforts of network engineers and operators to limit loss, it will probably never be eliminated due to the intrinsic dynamics and scaling properties of traffic in packet switched network.

Network operators have the ability to passively monitor nodes within their network for packet loss on routers using SNMP. End-to-end active measurements using probes provide an equally valuable perspective since they indicate the conditions that application traffic is experiencing on those paths [1], [2].

The rest of the paper is organized as section 2: discuss about the related work, section 3: presents the proposed model, section 4: discuss about the experimental setup, section 5: concludes the paper.

II. RELATED WORK

J. Bolot [9] and V. Paxson [10] have proposed algorithms to measure packet loss in the end-to-end networks. Yajnik [8] has evaluated packet loss correlations on longer time scales and developed Markov models for temporal dependence structures. Zhang [11] has characterized the packet loss in to several aspects based on their behavior. Papagiannaki [12] used a sophisticated passive monitoring infrastructure inside Sprint's IP backbone to gather packet traces and analyze characteristics of delay and congestion. Later, Sommers and Barford have discussed some of the drawbacks in standard end-to-end Poisson probing tools by comparing the loss rates measured by such tools to loss rates measured by passive means in a fully instrumented wide area infrastructure. The foundation for the notion that Poisson Arrivals See Time Averages (PASTA) was developed by Brumelle [13], and later formalized by Wolff [7]. Adaptation of those queuing theory ideas into a network probe context to measure loss and delay characteristic began with Bolot's study [9] and was extended by Paxson [7]. Baccelli [14] has analyzed the usefulness of PASTA in the networking. Several studies include the use of loss measurements to estimate network properties such as bottleneck buffer size and cross traffic intensity [15], [16].

The Internet Performance Measurement and Analysis efforts [17], [18] resulted in a series of RFCs that specify how packet loss measurements should be conducted.

However, those RFCs are devoid of details on how to tune probe processes and how to interpret the resulting measurements. ZING tool [19] is used for measuring end-to-end packet loss in one direction between two participating end hosts. It sends UDP packets at Poisson-modulated intervals with fixed mean rate. Savage [20] has presented STING tool to measure loss rates in both forward and reverse directions from a single host. This tool uses a clever scheme for manipulating a TCP stream to measure loss. Allman [21] has presented a method to estimate TCP loss rates from passive packet traces of TCP transfers taken close to the sender. Finally, M. Coates and N. Duffield [22, 23] have presented network tomography based on using both multicast and unicast probes for inferring loss rates on internal links on end-to-end paths.

III. PROPOSED MODEL

In our proposed model, we consider a tree-structured network consisting of single source and multiple receivers which is described in figure 1. From the source a distinct path is associated for every receiver. Each path consists of one or more links between nodes. If subpaths consisting of two or more links with no branches exist in the network, then those subpaths either removed or replaced by a single composite link. For packet transmissions, we assume a simple Bernoulli loss model for each link. The unconditional success probability of link i is defined as

$$\alpha_i \equiv \Pr(\text{packet successfully transmitted from } p(i) \text{ to } i),$$

where $p(i)$ denotes the index of the parent node of node i . A packet is successfully sent from $p(i)$ to i with probability α_i and may be dropped with probability $1-\alpha_i$. The loss processes of links are mutually independent of each other. Although spatial dependence may be observed in networks due to common traffic, such dependence is highly circumstantial and cannot be readily incorporated in a model that is intended to be generally applicable to a variety of networks.

Bolot [10] proposed a Markovian model of packet loss based on traffic in the internet. V. Paxson [7] has discussed that Markovian model do not fully account for the extended loss bursts. In our proposed model, we adopt a similar method for modeling the packet loss processes on each link. If two, back-to-back packets are sent from node $p(i)$ to node i , then the conditional success probability is defined as

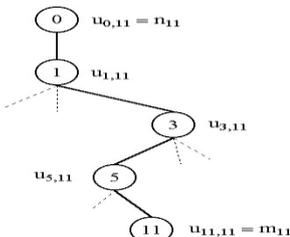


Figure 1: Tree structured network model

$$\beta_i \equiv \Pr(\text{2nd packet } p(i) \rightarrow i \mid \text{1st packet } p(i) \rightarrow i), \quad (1)$$

Where $p(i) \rightarrow i$ denotes the successful transmission of a packet from $p(i)$ to i . That is, given that the first packet of the pair is received, then the second packet is received with probability β_i and dropped with probability $1-\beta_i$. We anticipate that $\beta_i > \alpha_i$ for each i since knowledge that the first

packet was successfully received suggests that the queue for link i is not full.

Each link in the tree has unconditional and conditional success probabilities, α_i and β_i respectively. These probabilities will effect the measurement of packet loss in the end-to-end network. The packet loss can be measured in various ways such as UDP can be used for active probing or TCP connections may be passively monitored, in which back-to-back events are selected from the TCP traffic flows.

Single Packet Measurement: Suppose that n_i packets are sent to receiver i in that m_i packets are actually received and n_i-m_i are dropped. The likelihood of m_i given n_i is binomial and is given by

$$l(m_i | n_i, p_i) = \binom{n_i}{m_i} p_i^{m_i} (1-p_i)^{n_i-m_i} \quad (2)$$

Back-to-Back Packet Pair Measurement: Suppose that the source sends a large number of back-to-back packet pairs in which the first packet is destined for receiver i and the second for receiver j . We assume that the timing between pairs of packets is considerably larger than the timing between two packets in each pair. Let $n_{i,j}$ denote the number of pairs for which the first packet is successfully received at node i , and let $m_{i,j}$ denote the number of pairs for which both the first and second packets are received at their destinations. Furthermore, let $k_{i,j}$ denote the node at which the paths $p(0,i)$ and $p(0,j)$ diverge, so that $p(0,k_{i,j})$ is their common subpath. The $m_{i,j}$ likelihood of given $n_{i,j}$ is binomial and is given by

$$l(m_{i,j} | n_{i,j}, p_{i,j}) = \binom{n_{i,j}}{m_{i,j}} p_{i,j}^{m_{i,j}} (1-p_{i,j})^{n_{i,j}-m_{i,j}}$$

Where

$$p_{i,j} = \prod_{q \in \mathcal{P}(0,k_{i,j})} \beta_q \prod_{r \in \mathcal{P}(k_{i,j},i)} \alpha_r \quad (3)$$

IV. EXPERIMENTAL SETUP

Let us now consider the simple two-receiver network shown in Figure 2. Assume that we have made measurements of single packet and back-to-back packet

$$\mathcal{M} = \{m_i\}_{i=2,3} \cup \{m_{i,j}\}_{i,j=2,3}$$

$$\mathcal{N} = \{n_i\}_{i=2,3} \cup \{n_{i,j}\}_{i,j=2,3}$$

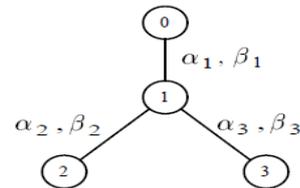


Figure 2: small network with two receivers

Maximum likelihood estimates of $\alpha_1, \alpha_2, \alpha_3$ are given by

$$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3) = \arg \max_{\alpha_1, \alpha_2, \alpha_3} \left[\max_{\beta_1, \beta_2, \beta_3} l(\mathcal{M} | \mathcal{N}, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) \right]$$

Note that direct optimization requires the joint maximization of the six dimensional likelihood function; a daunting task even in this simple case. Using the Expectation-Maximization (EM) Algorithm [15] we can

easily determine maximum likelihood in $O(k)$ time. where k is the number of iterations of the algorithm. For example

$$(\alpha_1, \alpha_2, \alpha_3) = (0.80, 0.90, 0.70)$$

$$(\beta_1, \beta_2, \beta_3) = (0.99, 0.99, 0.99)$$

$$\mathcal{N} = \{n_i = 10000\}_{i=2,3} \cup \{n_{i,j} = 10000\}_{i,j=2,3}$$

To evaluate the performance of proposed algorithm, we have implemented our proposed algorithm in NS2, which has been highly validated by the networking research community. The simulation parameters were listed in table 1.

Table 1: NS2 parameters

Parameters	Value
MAC Layer	IEEE 802.11
Number of nodes	20
Data rate	11Mbps
Packet Size	512 B
Simulation Duration	200 sec
Traffic Flow	TCP

Figure 3 and 4 shows results for the constant bit rate traffic with loss episodes of uniform duration.

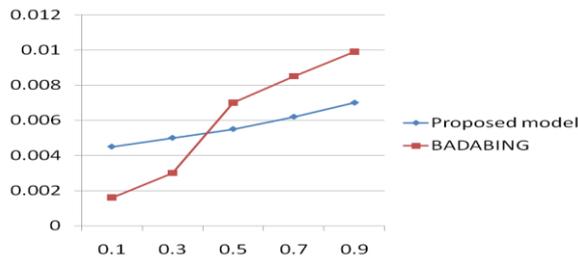


Fig 3: Packet loss frequency

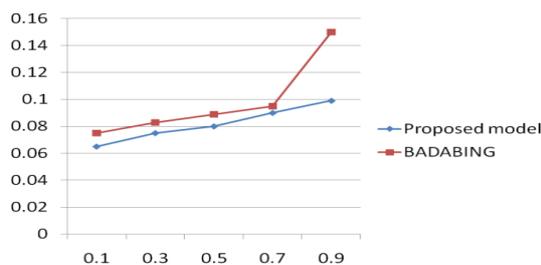


Fig 4: packet loss duration

V. CONCLUSION

Estimation of end-to-end packet loss in the network is very useful in various safety and non safety applications. Various tools are available in market to measure packet loss in the end-to-end networks. But all the existing tools are not able to calculate packet loss in the network accurately. In this paper we presented an algorithm for measuring end-to-end packet loss in the network. To evaluate the performance of the proposed algorithm, we simulate the algorithm in the NS2 simulator. The experimental results reveal that our proposed algorithm is very efficient in estimation of packet loss frequency and duration when compared BADABING tool.

REFERENCES

- [1] R. Cáceres, N. Duffield, J. Horowitz, and D. Towsley, "Multicast-based inference of network-internal loss characteristics," *IEEE Trans. Info. Theory*, vol. 45, November 1999.
- [2] C. Tebaldi and M. West, "Bayesian inference on network traffic using link count data (with discussion)," *J. Amer. Stat. Assoc.*, June 1998.
- [3] S. Vander Wiel, J. Cao, D. Davis, and B. Yu, "Time-varying network tomography: router link data," in *Proc. Symposium on the Interface: Computing Science and Statistics*, (Schaumburg, IL), June 1999.
- [4] Y. Vardi, "Network tomography: estimating sourcedestination traffic intensities from link data," *J. Amer. Stat. Assoc.*, 1996, pp. 365–377.
- [5] J.-C. Bolot, "End-to-end packet delay and loss behaviour in the Internet," in *Proc. SIGCOMM '93*, pp. 289–298, Sept. 1993.
- [6] J. O. Berger, B. Lisco, and R. L. Wolpert, "Integrated likelihood methods for eliminating nuisance parameters," *Work Paper 97-01*, Institute of Statistics & Decision Sciences, vol. Duke University, 1997, pp. Durham, NC.
- [7] R. Wolff, "Poisson arrivals see time averages," *Oper. Res.*, vol. 30, no. 2, Mar.–Apr. 1982.
- [8] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, "On the constancy of internet path properties," in *Proc. ACM SIGCOMM Internet Measurement Workshop '01*, San Francisco, CA, Nov. 2001.
- [9] J. Bolot, "End-to-end packet delay and loss behavior in the internet," in *Proc. ACM SIGCOMM '93*, San Francisco, CA, Sep. 1993.
- [10] V. Paxson, "End-to-end internet packet dynamics," in *Proc. ACM SIGCOMM '97*, Cannes, France, Sep. 1997.
- [11] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and modeling of temporal dependence in packet loss," in *Proc. IEEE INFOCOM '99*, New York, Mar. 1999.
- [12] D. Papagiannaki, R. Cruz, and C. Diot, "Network performance monitoring at small time scales," in *Proc. ACM SIGCOMM '03*, Miami, FL, Oct. 2003.
- [13] S. Brumelle, "On the relationship between customer and time averages in queues," *J. Appl. Probabil.*, vol. 8, 1971.
- [14] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot, "The role of PASTA in network measurement," in *Proc. ACM SIGCOMM*, Pisa, Italy, Sep. 2006.
- [15] S. Alouf, P. Nain, and D. Towsley, "Inferring network characteristics via moment-based estimators," in *Proc. IEEE INFOCOM '01*, Anchorage, AK, Apr. 2001.
- [16] K. Salamatian, B. Baynat, and T. Bugnagel, "Cross traffic estimation by loss process analysis," in *Proc. ITC Specialist Seminar on Internet Traffic Engineering and Traffic Management*, Wurzburg, Germany, Jul. 2003.
- [17] Merit Internet Performance Measurement and Analysis Project. 1998 [Online]. Available: <http://www.nic.merit.edu/ipma/>
- [18] Internet Protocol Performance Metrics, <http://www.advanced.org/IPPM/index.html>
- [19] A. Adams, J. Mahdavi, M. Mathis, and V. Paxson, "Creating a scalable architecture for Internet measurement," *IEEE Network*, 1998.
- [20] S. Savage, "Sting: A tool for measuring one way packet loss," in *Proc. IEEE INFOCOM '00*, Tel Aviv, Israel, Apr. 2000.
- [21] M. Allman, W. Eddy, and S. Ostermann, "Estimating loss rates with TCP," *ACM Perform. Eval. Rev.*, vol. 31, no. 3, Dec. 2003.
- [22] M. Coates and R. Nowak, "Network loss inference using unicast end-to-end measurement," in *Proc. ITC Conf. IP Traffic, Measurement and Modeling*, Sep. 2000.
- [23] N. Duffield, F. Lo Presti, V. Paxson, and D. Towsley, "Inferring link loss using striped unicast probes," in *Proc. IEEE INFOCOM '01*, Anchorage, AK, Apr. 2001.

Market power Analysis Based on Relaxation Algorithm and the Nikaido-Isoda Function in Electricity Markets

Davoud Abbasi¹, Mohsen Assili², Seyed Alireza Razavian³

* Department of Electrical Engineering, Shahrood University of technology, Iran

** Department of Electrical Engineering, Islamic Azad University, Shahrood branch, Iran

Abstract: The issue of evaluating market power is an important challenge in power system planning. This paper develops an analysis based on a relaxation algorithm and Nikaido–Isoda function for calculation of Nash equilibrium and evolution of market structure and performance. This is done through the development of a generation company trading game that, via Nikaido–Isoda function, simulates how players coordinate their behavior in generation to maximize their profit. In this paper, with changing type of power plants belonging to generation company and demand elasticity, market performance is evaluated. The results for pool-based markets shows, generation companies coordination is a very important factor and able to decrease competition and efficiency in market.

Keywords: Market power, Nikaido–Isoda, relaxation algorithm, Nash equilibrium, electricity markets, demand elasticity

I. INTRODUCTION

All competitive markets are free markets, but not all free markets are competitive. Markets where one or more generation companies (Gen Co) have the ability to raise price and profit are not perfectly competitive.

All capital intensive industries manifest a co-evolution of market structure and performance, but because of the instantaneous, non-storable nature of electricity, low demand elasticity, high requirements for security of supply and wide seasonal variations this co-evolution is not deterministic. This means that electricity is provided from an economic and technical mix of base load, mid-merit and peaking plant is not a certain parameter. This raises the strategic issue for competing companies to evolve towards with a mix of different kinds of generations or which players are more dominant in the base, mid or peaking segments of the market [1]. The traditional assessment of market power has focused on the supplier's ability to profitably alter prices away from competitive levels [2] and how major players manage its contribution of peak, mid and base load plant, in order to set market prices with their marginal plant and thereby reap higher profit contributions [1]. [3] suggested in the liberalized markets, different segments would emerge at least for base load and peak plants. [4] is looked at generator bid and cost data to analyze market power for two largest generating companies in the England and Wales electricity pool.

Capacity withholding is analyzed as an important parameter in market power in [5] and has shown electric generating firms whose market shares range between 10 percent and 40 percent may be profitable.

In [6] this mechanism is used as an effective way to exercise market power in the electricity spot market of England and Wales.

II. DEFINITIONS AND CONCEPTS

1.1 Nikaido-Isoda Methodology

An N-person game can be used as a mathematical model of electricity market. In this game a number of players (electricity companies) interact in a setting of strategies. This means that the profit of each player depends on his own actions and on the actions of the other participants in the game.

Let N be the number of players. The i_{th} player has a set of strategies X_i and ϕ_i is the profit function of i_{th} player. The collective action set $X = X_1 \times X_2 \times \dots \times X_N$ be the vector formed by all these decision variables. Each player makes a choice according to his own strategy for maximizing its profit. In real electricity market each Gen Co (player) able to have enough information about its own and other players past actions. This is called the information set.

Assume that there are players participating in a game. Each player can take an individual action represented by vector x_i . All players, when acting together, can take a vector $x = x_1 \times x_2 \times \dots \times x_N$ that is a subset of X.

The vector x is defined as the joint action vector [7] formed by the strategies of each player. Also

$\left(\frac{y_i}{x} \right) = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)$ is defined as the

vector of strategies that player i can take, while the strategies of the other players $x_j, j \in \{1, 2, \dots, i-1, i+1, \dots, N\}$ remaining constant.

Nash equilibrium point can be expressed as $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ if for each i it is true that

$$\phi_i(x^*) = \max_{(x_i|x^*) \in X} \phi_i(x_i|x^*) \quad (1)$$

Notice that at x^* no player can improve his individual profit by his own action. For finding Nash equilibrium point a general way is using the Nikaido–Isoda function. The Nikaido-Isoda function is defined as [8]

$$\Psi_{(x,y)} = \sum_{i=1}^N \left[\phi_i \left(\frac{y_i}{x} \right) - \phi_i(x) \right] \quad (2)$$

It follows from the definition of the Nikaido–Isoda function that in Nash point $\Psi_{(x,y)} = 0$. Each summand of the Nikaido–Isoda function can be thought of as the change in the profit of a player when his action changes from x_i to y_i while all other players continue to play according to x. The function thus represents the sum of these changes in profit functions. Maximum value that this function can take by changing y, for a given x, is always greater than zero except in Nash equilibrium point. At equilibrium point, no player can make a unilateral improvement to their profit, and so in

this case maximum value of Nikaido–Isoda function can be zero.

In conclusion, when the Nikaido–Isoda function satisfies certain concavity conditions [7] and cannot be made positive for a given y , the Nash equilibrium point is reached. This is used to construct a termination condition for the relaxation algorithm that explained in next section.

Finally, the optimum response function at point x can be defined as

$$Z(x) = \arg \max_{y \in X} \Psi(x, y), x, Z(x) \in X \quad (3)$$

It is the result of maximizing the Nikaido–Isoda function, where all players try to improve their profits. This function returns the set of players’ actions whereby they all try to unilaterally maximize their respective profits with respect to actions y_i and so, by “playing” actions $Z(x)$ rather than x , the players approach the equilibrium [9]. Note that, by doing that, a player maximizes its profit assuming that the competitors are fixed in their actions, which is the definition of Nash-Cournot equilibrium. In the next section, a relaxation algorithm that uses the Nikaido–Isoda function to compute Nash equilibrium is presented. In simple Nikaido–Isoda function the players wish to move to a point that represents an improvement in compare with current player situations and these movements may cause system instability. The relaxation algorithm adds a ratio of previous value of x to the new response $Z(x)$. Technical definitions that are used in the convergence theorem of the algorithm are expressed in [7].

1.2 Relaxation Algorithm

In order to find a Nash equilibrium of a game, having an initial estimate x_0 , the relaxation algorithm of the optimum response function, when $Z(x)$ is single-valued (every input is associated with one output only) and the concavity conditions are satisfied (the Nikaido-Isoda function is weakly convex-concave [8]) is

$$x(j+1) = (1-\alpha(j))x(j) + \alpha(j)Z(x(j)) \quad j=0,1,2,\dots \quad (4)$$

where $0 < \alpha(j) < 1$. An iterative algorithm is constructed as a convex combination of the improvement point $Z(x(j))$ and the current point $x(j)$. The optimum response function $Z(x(j))$ is calculated after solving an optimization problem as seen in (3). This averaging ensures convergence of the algorithm under certain conditions [7].

It is interesting to note that we can consider the algorithm as either performing a static optimization or calculating successive actions of players in convergence to equilibrium in a real time process. If all profits are known to us, we can directly find the Nash equilibrium using the relaxation algorithm. However, if we only have access to one player’s profit function and all players past actions, then at each stage in the real time process the optimum response should be chosen for that player, assuming that the other players will play as they had in the previous period. In this way, convergence to the Nash normalized equilibrium will occur at $j \rightarrow \infty$

Thus, by taking a sufficient number of iterations, the algorithm converges to the Nash equilibrium. The problem can be either considered a centralized optimization model or a calculation of the succession of actions by the players at each stage, where players choose their optimum response

given the actions of the opponents in the previous period. The theorem that ensures convergence of the relaxation algorithm is presented in full detail in [7].

$\alpha(j)$ is an important factor to converge and optimize the convergence rate of algorithm. In [7] $\alpha(j)$ is supposed constant until convergence conditions are reached, and thereafter decaying with factors $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$. We have found that using a constant value lower than 0.4 leads to a good and approximately quick convergence but values higher than 0.4 in most of our experiments leads to algorithm divergence. At last suitable $\alpha(j)$ values in any applications may be obtained by trial and error.

1.3 HHI index

The most widely used measure of concentration in a market is the Herfindahl–Hirschman Index (HHI). HHI is linked directly to market power in one theoretical model of competition [10]. This parameter widely applied in competition law and also technology management. It is defined as the sum of the squares of the market shares, where the market shares are expressed as fractions. The result is proportional to the average market share, weighted by market share.

$$HHI = \sum S_i^2 \quad (5)$$

Where S_i is market share of each generation company [10].

If the market share of each company is expressed in percentage terms, the HHI lies between 0 and 10,000. The maximum value of the HHI occurs when market is monopoly and without any competition. The minimum value of the HHI occurs in the limit that the generation companies comprise a very large number of companies, each with negligible market shares.

Increases in the HHI index generally indicate a decrease in competition and an increase of market power, whereas decreases indicate the opposite. The major benefit of the HHI is in relationship to such measures as the concentration ratio is that it gives more weight to larger companies.

III. CASE STUDIES

The most sensible method of calculating market power impacts in an electricity market is to simulate the operation of that electricity market and, thereby, directly measure the price and revenue impacts of generation company strategies.

The methodology presented above is used to simulate the market using Nash-Cournot equilibrium computed with Nikaido-Isoda function and relaxation algorithm, of some case studies, assuming a pool-type market.

In all case studies, it is assumed that there are three generating companies and each of them possesses several generating unit, as shown in Table I. Proposed model has 19 units. The cost of a generating unit is defined as:

$$C_i(P_{gi}) = \frac{C_i}{2} P_{gi}^2 + d_i P_{gi} + e_i \quad (6)$$

C_i , d_i and e_i coefficients are shown in table 1 and P_{gi} is generation of unit.

Table 1: system data of proposed model

No of unit	Case A Company type	Case B Company type	Case C Company type	Case D Company type	Minimum power (MW)	Maximum power (MW)	C_i (\$/MW ² h)	d_i (\$/MWh)	e_i (\$/h)
1	1	1	1	1	0	736	0.00021	0	0
2	2	1	1	1	0	501.8	0.0003	0	0
3	3	1	1	1	0	354.8	0.00046	0	0
4	3	1	1	2	0	261	0.0006	0	0
5	1	1	2	1	0	340	0.00466	31.68	0
6	1	1	3	1	0	379	0.00194	14.68	0
7	2	1	2	1	0	379	0.00171	12.97	0
8	3	1	3	1	0	368.4	0.00183	13.52	0
9	2	1	2	2	0	304	0.00377	22.91	0
10	3	1	3	3	0	250	0.00471	23.57	0
11	3	1	3	3	0	244.9	0.00347	16.99	0
12	1	1	2	2	0	128	0.01099	28.13	0
13	2	1	3	3	0	108	0.04202	90.76	0
14	3	1	2	2	0	97	0.06649	129	0
15	3	1	3	3	0	58	0.11197	129.88	0
16	2	1	2	2	0	49	0.07562	74.11	0
17	1	1	3	3	0	23.8	0.26658	126.89	0
18	1	1	2	2	0	16	0.05076	58.88	0
19	3	1	3	3	0	12	0.41042	98.5	0

Units with number 1 to 4 are baseload power plants and have low cost coefficients in compare with other units. Units with numbers 5 to 11 have a few higher cost coefficients and is classified as mid merit units. Remaining units are peak.

Other details of each power plant including minimum power, maximum power and owner of unit in each case study are given in table I.

In case A, all Gen Cos have approximately similar portion of baseload, mid merit and peak units. In case B, the market is monopoly and one company is owner of total units and for case C, all baseload units are owned by company no 1 and mid merit and peak units are divided equally between other generation companies.

In D, company no 1 is dominant in baseload and mid merit units and other companies have only a few portion of units in this two parts.

To determine each company's profit function ϕ_i as used in (1), price is assumed a strictly decreasing function of the electricity demand [7].

$$p = \alpha - \beta P_{load} \quad (7)$$

The income function of each company is equal to its generation multiplied by price. Now profit function of each company is defined as [7]:

$$\phi_i(P_{gi}) = pP_{gi} - C_i(P_{gi}) = pP_{gi} - \frac{C_i}{2} P_{gi}^2 + d_i P_{gi} + e_i \quad (8)$$

That should be computed for all units.

All case studies, with above assumption are simulated in three different demand elasticity for covering baseload, mid merit and peak zones and nash-cournot equilibrium is found with combination of relaxation algorithm and nikaido-isoda function. Final results are presented in table 2 to 9.

Table 2: Case A system Nash equilibrium results

	Company no 1	Company no 2	Company no 3	Total profit	Price
profit Case 1	516448.72	508797.93	513686.67	1538933.32	1023.4
profit Case 2	1264799.64	1214044.41	1241125.49	3719969.54	1673.89
profit Case 3	4389293.47	3702676.98	4479941.43	12571911.88	3229.17

Table 3: the generation details for Case A and HHI index

	Total	Company no 1	Company no 2	Company no 3	HHI
Generation Case 1	1513.3	33.52%	33.1%	33.38%	3333.43
Generation Case 2	2268.05	33.56%	33.06%	33.37%	3333.46
Generation Case 3	4055.41	34.9%	29.51%	35.59%	3355.51

Table 2 shows Nash equilibrium results for case A (balanced market). Case 1 is for elasticity in base load, case 2 for mid merit and case 3 for peak zone. Corresponding to each zone, the price and generation is increased in table 2.

Table 3 shows that HHI index is at its minimum value for 3 company and portion of companies in market is approximately constant in all demand situations and market is clearly competitive.

Table 4: Case B system Nash equilibrium results

	Company no 1	Total profit	Price
profit Case 1	2046027.02	2046027.02	2048.82
profit Case 2	4810436.86	4810436.86	3141.53
profit Case 3	15902321.57	15902321.57	5785.97

Table 5: the generation details for Case B and HHI index

	Total	Company no 1	HHI
Generation Case 1	1000.59	100	10000
Generation Case 2	1534.23	100	10000
Generation Case 3	2777.01	100	10000

Table 4 shows Nash equilibrium results for case b (monopoly market). Similar to case A, Corresponding to each zone, the price and generation is increased in table 4.

It is obvious, that price and total profit in each case is higher than table 2 (about 100%), because the market is monopoly and not competitive. In fact in a monopoly market, market owner (dominant generation company) constrains its strategy to customers and set market price to its marginal price and for this reason will obtain maximum profit. Table 5 shows that HHI index is at its maximum value and company no 1, with a generation about 30% lower, earn a higher profit than case A.

Table 6: Case C system Nash equilibrium results

	Company no 1	Company no 2	Company no 3	Total profit	Price
profit Case 1	620793.95	415518.76	453096.35	1489409.06	1122.1
profit Case 2	1407778.09	1005773.82	1121793.74	3535345.65	1686.11
profit Case 3	5264842.32	3488842.61	3772183.04	12525867.97	3248.43

Table 7: the generation details for Case C and HHI index

	Total	Company no 1	Company no 2	Company no 3	HHI
Generation Case 1	1463.95	37.87	29.89	32.24	3366.95
Generation Case 2	2261.94	36.99	30.24	32.78	3356.6
Generation Case 3	4045.78	40.14	29.21	30.65	3403.94

Nash equilibrium results for case C are given table 6. In this case, company 1 has all base load plants and market power belongs to this company. Company 1 has higher generation in all zones and due to its low cost plants, earns more profits than others. Price in this case is a few higher than fully competitive market but is so lower than monopoly market. Table 7 shows that market concentration is a few higher than case A. generation details given in table 7 apparently explain company 1 imposed its generation strategy to market.

Nash equilibrium and HHI results for case D are given table 8 and 9. In this case, company 1 has a large number of baseload and mid merit plants and market power belongs to this company more than case C.

Table 8: Case D system Nash equilibrium results

	Company no 1	Company no 2	Company no 3	Total profit	Price
profit Case 1	674989.85	464005.19	391417.16	1530412.2	1166.86
profit Case 2	1827767.46	1193932.73	826664.13	3848364.32	1917.45
profit Case 3	8403715.89	3242221.7	2529087.17	14175024.76	4197.44

Table 9: the generation details for Case D and HHI index

	Total	Company no 1	Company no 2	Company no 3	HHI
Generation Case 1	1441.57	40.22	30.49	29.29	3405.3
Generation Case 2	2146.27	44.52	32.26	23.22	3562.02
Generation Case 3	3571.28	56.57	23.93	19.5	4152.95

Price is increased in all demand situations in compare with case A and C. Company 1 has more generation in all zones and due to its plants, earns more profits than others. Table 9 shows that market concentration is higher than case A and C. In this case company 1 has the ability to raise market price and its profit due to its portion of generation.

IV. CONCLUSION

The issue of evaluating market power is an important challenge in restructured electricity markets. The model presented in this paper carries out an iterative Nash-Cournot equilibrium game that considers a generating pool market model in some practical case to evaluate market power. For solving game, an algorithm based on the Nikaido–Isoda function and a relaxation algorithm is used. It allows for the incorporation of the network constraints and different type of plants. It is possible to add independent demand curves for each time and thereby assigning different elasticity and consumption values and evaluate market power separately. To evaluate market power, HHI index as an economic parameter for showing concentration in markets is used. Our model is a useful tool to analyze the market power and strategic behavior of the agents in a competitive electricity market and can be used by independent system operator to enhance market efficiency.

Four case studies of electricity markets are presented. The first case study shows a balanced market. In this case HHI index is near to ideal value for all demand situations. In case B, a monopoly market is simulated and results, showed a very high price increases in compare with case A. HHI index in this situation is at its maximum value. In case C and D, Changes in generation company properties increase market power and decrease competition in market. Results shows that in different elasticity, different sets of coalitions among generating companies can be change significantly market power. As a final conclusion the results for pool based markets shows, generation companies coordination according to demand elasticity is a very important factor and able to decrease competition and efficiency in market.

REFERENCES

- [1] F. Oliveira, D. W. Bunn Agent-Based Analysis of Technological Diversification and Specialisation in Electricity Markets, *European Journal of Operational Research*, Jan 2007, Issue 181, p. 1265-1278
- [2] Stoft, S. *Power System Economics: Designing Markets for Electricity*, IEEE Press, 2002.
- [3] Borenstein, S., J. Bushnell, E. Kahn, And S. Stoft, Market power in California electricity markets, *Utilities Policy*, 1995, 5 (3-4), 219-236.
- [4] Von der Fehr, N-H and Harbord, D. Spot Market Competition in the UK Electricity Industry, *The Economic Journal*, 1993, Vol. 103, pp. 531 – 546.
- [5] Newbery, D.M., Regulatory Policies and Reform in the Electricity Supply Industry, *Cambridge Working*

Papers in Economics 9421, 1995, Faculty of Economics, University of Cambridge.

- [6] Wolak, F.A., Patrick, R.H. Industry structure and regulation in the England and Wales electricity market. 1997. *Crew, M. (Ed.), Pricing and Regulatory Innovation under Increasing Competition. Kluwer Academic Publishers, Dordrecht, Netherlands.*
- [7] J. Contreras, M. Klusch, and J. B. Krawczyk, Numerical solutions to Nash-Cournot equilibria in coupled constraint electricity markets, *IEEE Trans. Power Syst.* , Feb. 2004, vol. 19, no. 1, pp. 195–206,.
- [8] J. B. Krawczyk and S. Uryasev, Relaxation algorithms to find Nash equilibria with economic applications, *Environment. Model Assess. Jan. 2000, vol. 5, no. 1*, pp. 63–73, . [Online]. Available: http://www.ise.ufl.edu/uryasev/relaxation_algorithms_for_Nash_equilibrium.pdf.
- [9] Molina, J.P. ,Zolezzi, J.M. , Contreras, J. , Rudnick, H. , Reveco, M.J. Nash-Cournot Equilibria in Hydrothermal Electricity Markets , *IEEE Trans. Power Syst.* , Aug. 2011, vol. 26, no. 3, pp. 1089-1101.
- [10] Williams, Edward and Richard Rosen., A Better Approach to Market Power Analysis, 1999, *Working Paper*, available at <http://www.tellus.org/general/publications.html>.

Recent Advancement In Electric Discharge Machining, A Review

Amandeep singh¹, Neel kanth grover², Rakesh sharma³

¹(Department of Mechanical Engg department FCET Ferozpur)

²(Department of Mechanical Engg Department SBSCET Ferozpur)

ABSTRACT: Electric discharge machining is non conventional machining process used for machining of hard materials which cannot machined by conventional machining process. Electric discharge machining is an electro sparking method of metal working involving an electric erosion effect. A pulse discharge occurs in a small gap between the work piece and the electrode and removes the unwanted material from the parent metal through melting and vaporizing

Keyword: Cryogenic, EDM, surface finish, Wear

I. INTRODUCTION

Electric discharge machining is a non conventional machining process and has found its wide application in making moulds, dies, and in aerospace products and in surgical equipments[1]. The process is based on removing material from a part by means of a series of repeated electrical discharges between tool called the electrode and the work piece in the presence of a dielectric fluid [2]. The electrode is moved toward the work piece until the gap is small enough so that the impressed voltage is great enough to ionize the dielectric [3]. Short duration discharges are generated in a liquid dielectric gap, which separates tool and work piece. The material is removed with the erosive effect of the electrical discharges from tool and work piece [4]. EDM machining does not involves direct contact of tool and work piece [1].Material of any hardness can be cut by EDM only condition it should be electrically conductive [5]. In this review paper, various trends in electric discharge machining has been considered involving powder mixed electrolyte used for EDM, incorporating tool vibration, green EDM (dry EDM), treatment of electrode used for EDM, and validating EDM performance using modeling techniques.

II. POWDER MIXED ELECTROLYTES.

In this process the fine abrasive particles are mixed with the electrolyte. Powder mixed EDM process is called hybrid machining process and called as PHEDM [6]. Electrically conductive powder reduces the insulating strength of the dielectric fluid and increases the spark gap between the tool and the work piece. EDM process becomes more stable and improves machining efficiency, MRR and surface quality.However; most studies were conducted to evaluate the surface finish since the process can provide mirror surface finish which is a challenging issue in EDM. The characteristics of the powder such as the size, type and concentration influence the dielectric performance [7].

2.1 SURFACE FINISH

Ming and He [8] indicates that some conductive powder can lower the surface roughness and the tendency

of cracks in middle finish and finish machining but the inorganic oxide additive does not have such effect. Wong et al. [9] compares the near mirror- finish phenomenon using graphite, silicon (Si), aluminum (Al), crushed glass, silicon carbide (SiC) and molybdenum sulphide with different grain size. Al powder has been reported to give mirror finish for SKH-51 work pieces, but not on SKH-54 work pieces. They suggested that it is important to have the correct combination of Powder and work piece materials. Fu-chen [10] investigates the effect of powder properties on SQ of SKD-11 work piece using Al, chromium (Cr), Copper (Cu), and SiC powders. The smallest particle (70–80 nm) generates best surface finish and Al powder produces the best surface finish.

2.2 MATERIAL REMOVAL RATE

Jeswani [11] revealed that the addition of about 4 g/l of Fine graphite powder in kerosene increases MRR by 60% and tool wear by 15%. Yan and Chen [12] describes the effect of dielectric mixed with electrically conductive powder such as Al powder on the gap distance, surface roughness, material removal rate, relative electrode wear ratio, and voltage. waveform. It is shown that the dielectric with suspended electrically conductive powder can enlarge the gap distance and can improve the energy dispersion, surface roughness, and material removal rate. Machining efficiency and surface roughness of rough PMEDM in rough machining was studied by Zhao et al.[12] using Al with 40 g/l and 10 mm granularity and they discovered that machining efficiency was improved from 2.06 to 3.4mm³/min with an increase in rate of 70%. The Machining efficiency can be highly increased by selecting Proper discharge parameter (increasing peak current and Reducing pulse width with better surface finish in comparison with conventional EDM machining. Y.F.

Tzeng, and Lee [13] indicated that the greatest MRR is produced by chromium and 70–80 nm of grain size. Kansal et al. [14] established optimum process conditions for PMEDM in the rough machining phase using the Taguchi method with Graphite powder and found out that addition of an appropriate amount of the graphite powder into the dielectric fluid caused discernible improvement in MRR and reduction in tool wear as well as in surface roughness. Pec [15] investigates that powder mixed electrolyte conditions promotes the reduction of surface roughness, crater diameter, crater depth and the white-layer thickness. it was confirmed that electrode area influence on the surface quality produced. This influence was mathematically described by several linear equations relating the surface roughness, the white-layer thickness, the crater depth and width to the electrode area as the independent variable. Moreover, it was found that the sensitivity of the surface quality measures to the electrode area is smaller when mixed-powder dielectric is used. Also, powder-mixed dielectric significantly reduces surface

heterogeneity contributing to increase process robustness.

So, it contributes to the performance of the EDM process particularly when large electrode areas are involved and when a high-quality surface is a requirement.

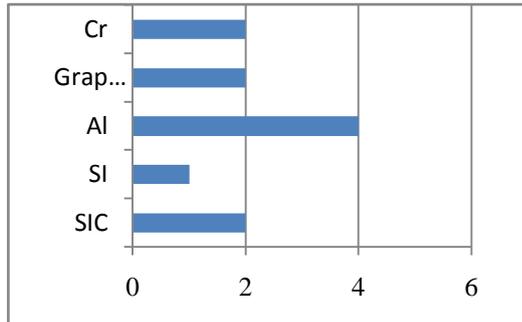


Fig. Various additives used in electrolyte based on collected papers

2.3 CONCLUSION

Use of powder in electrolyte provide mirror like surface finish and increase in material removal rate. Proper work piece and powder combination must be used for better results...

III. TOOL VIBRATION

Introduction of ultrasonic vibration to the electrode is one of the methods used to expand the application of EDM and to improve the machining performance on difficult to machine materials. The study of the effects on ultrasonic vibration of the electrode on EDM has been undertaken since mid 1980s. The higher efficiency gained by the Employment of ultrasonic vibration is mainly attributed to the better circulation of dielectric and debris removal from work piece [16]. Zhang et al. [17] proposed spark erosion with ultrasonic frequency using a DC power supply instead of the usual pulse power supply. The pulse discharge is produced by the relative motion between the tool and work piece simplifying the equipment and reducing its cost. They have indicated that it is easy to produce a combined technology which benefits from the virtues of ultrasonic machining and EDM.

3.1 VIBRATION, ROTARY AND VIBRO-ROTARY

Ghoreishi and Atkinson [18] compared the effects of high and low frequency forced axial vibration of the electrode, rotation of the electrode and combinations of the methods (vibro-rotary) in respect of MRR, tool wear ratio (TWR) and surface quality (SQ) in EDM die sinking and found that vibro-rotary increases MRR by up to 35% compared with vibration EDM and by up to 100% compared with rotary EDM in semi finishing.

3.2 CONCLUSION

Ultrasonic vibration makes the equipment simple and increases the material ejection from work piece.

IV. DRY MACHINING

In dry EDM, tool electrode is formed to be thin walled pipe. High-pressure gas or air is supplied through the pipe. The role of the gas is to remove the debris from the gap and to cool the inter electrode gap. The technique was developed to decrease the pollution caused by the use of liquid dielectric which leads to production of vapor

during machining and the cost to manage the waste. Yu et al. [19] investigated the capability of the technique in machining cemented carbide material and compared the machining characteristics between oil EDM milling and oil die sinking EDM. They found that for machining the same shape, oil die sinking EDM shows shorter machining time. But because oil die sinking requires time for producing Electrodes, dry EDM should be more useful in actual production. The information given in this paper is interesting and they are reproduced here for better clarity.

Figs. 1 and 2 show the work removal rate and electrode Wear ratio in groove machining. According to the results, work removal rate of dry EDM milling is about six times larger than that of oil EDM milling, and electrode wear ratio one-third lower. In Fig. 3, it is shown that the EDM method with the shortest machining time was oil die sinking EDM, dry EDM milling was second, and oil EDM milling third. The lowest electrode wear ratio machining was dry EDM milling (see Fig. 4).

4.1 CONCLUSION

Dry EDM is eco friendly machining. Pollution is reduced by use of gas instead of oil based dielectric. Work removal rate also get enhanced by dry EDM.

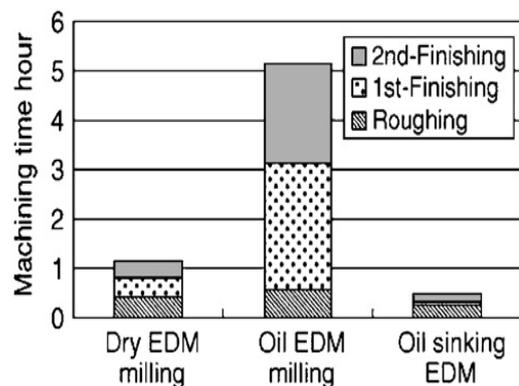
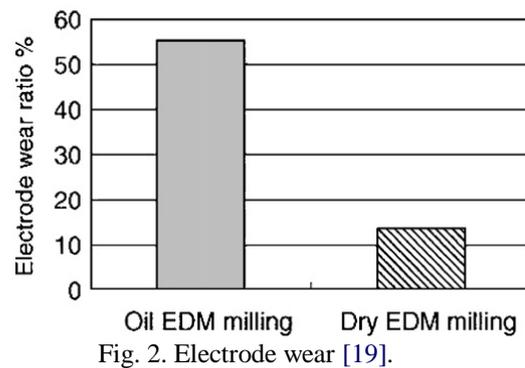
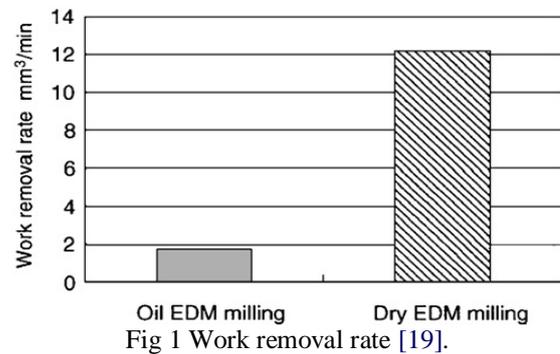


Fig 3 Machining time [19].

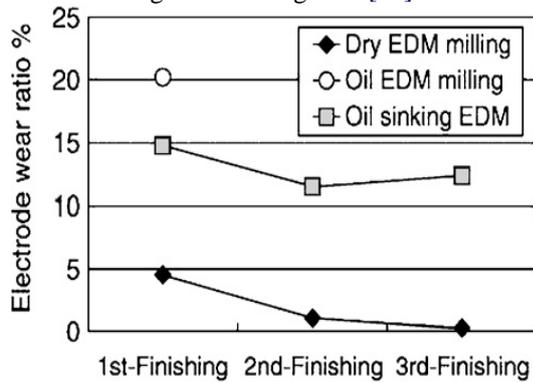


Fig. 4. Electrode wear ratio [19].

V. CRYOGENIC TREATMENT OF TOOL

The word cryogenic comes from the Greek word “kryos” which means cold. Cryogenic is the science of study of material at low temperature at which the properties of materials significantly change. Cryogenics processing is the treatment of the materials at very low temperature around 77K. This technique has been proven to be efficient in improving the physical and mechanical properties of the materials such as metals, plastics and composites. It improves the wear, abrasion, erosion and corrosion resistivity, durability and stabilizes the strength characteristics of various materials. Darwin [20] investigated that deep cryogenic treatment (DCT) is a one-time permanent process, carried out on steel components in such a way that the material is slowly cooled down to the cryogenic temperature, after which it is held at that temperature for a specified period of time and is heated back to room temperature at a slow rate followed by low temperature tempering. The DCT has a lot of benefits. It not only gives dimensional stability to the material, but also improves wear resistance, strength and hardness of the materials. Kalia [21] showed that cryogenic refines and stabilizes the crystal lattice structure and distribute carbon particles throughout the material resulting in a stronger and hence more durable material.

5.1 TOOL WEAR

Abdulkarim [22] investigated the effect of cryogenic cooling on electrode wear and surface roughness of the work piece used in EDM process. By using the cryogenic cooling the electrode wear ratio reduced up to 27% and there was 8% improvement in the surface roughness. MRR can be increased with decrease in electrode wear by controlling the machine parameters. Copper is the main electrode in the EDM process because it has low electrical resistance and which help to transfer effectively energy to the work piece. The cooling effect of liquid nitrogen improves the electrical and thermal conductivity of copper and this effect in efficient heat transfer from the Cu electrode resulting in reducing the electrode wear. Singh [23] proved that cryogenic treatment improves the tool wear rate by 58.77% & reduces the surface roughness by 8%. For cryogenic Ti work piece minimum value (0.00943443 gm/min) of tool wear rate was achieved with cryogenic Ti as tool and value of current was 2 ampere. for plain Ti work piece minimum value of TWR (0.010418803 gm/min) was achieved with plain as Ti as tool and value of the current was 2 ampere. M.

Dhananchezian [24] examined the machining of stainless steel inherently generates high cutting temperature, which not only reduces tool life but also impairs the work piece surface quality. Conventional cooling methods are ineffective in controlling the high cutting temperature and rapid tool wear. In the present research work, the effect of liquid nitrogen as a coolant applied through holes made on the rake and flank surfaces of the PVD TiAlN coated tungsten carbide turning tool inserts of ISO CNMG 120412 MP-KC5010 on the turning of AISI 304 stainless steel is studied. The influence of cryogenic cooling on the cutting temperature, cutting force, surface roughness, and tool wear, has been compared with those of wet machining. It has been observed that in the cryogenic cooling method, the cutting temperature was reduced by 44–51%, the cutting force was decreased to a maximum of 16%, and the surface roughness was reduced by 22–34% over that of wet machining. Cryogenic cooling using liquid nitrogen reduced tool wear through the control of temperature-dependant wear mechanisms.

5.2 CONCLUSION

Tool wear is reduced to large extent by cryogenic treatment done on electrode. Surface roughness produced also lowered by use of cryogenic treated electrodes

VI. EDM IN WATER

Water as dielectric is an alternative to hydrocarbon oil. The approach is taken to promote a better health and safe environment while working with EDM. This is because Hydrocarbon oil such as kerosene will decompose and release harmful vapour (CO and CH₄) [25]. Research over the last 25 years has involved the use of pure water and water with additives.

6.1 PURE WATER

The first paper about the usage of water as dielectric was Published by Jeswani [26] in 1981. He compared the Performances of kerosene and distilled water over the pulse energy range 72–288 mJ. Machining in distilled water resulted in a higher MRR and a lower wear ratio than in kerosene when a high pulse energy range was used. With distilled water, the machining accuracy was poor but the surface finish was better. Tariq Jilani and Pandey [27] investigated the performance of water as dielectric in EDM using distilled water, tap water and a mixture of 25% tap and 75% distilled water. The best machining rates have been achieved with the tap water and machining in water has the possibility of achieving zero electrode wear when using copper tools with negative polarities. Konig and Siebers [28] explained the influence of the working medium on the removal process. They indicated that working medium has a sustained influence on the removal process. The erosions process in water-based media consequently possesses higher thermal stability and much higher power input can be achieved especially under critical conditions, allowing much greater increase in the removal rate. A considerable difference between conventional oil based dielectrics and aqueous media in term of specific boiling energy of aqueous media is eight times higher than oil based dielectric and boiling phenomena occur at a lower temperature level. During investigating on white surface layer, Kruth et al. [29] found

that the use of an oil dielectric increases the carbon content in the white layer and appears as iron carbides (Fe_3C) in columnar, dendritic structures while machining in water causes a decarbonization. While investigating the influence of kerosene and distilled water as dielectric on Ti-6Al-4V work pieces. Chen [30] found that carbide is formed on the work piece surface while using kerosene while oxide is formed on the work piece surface while using distilled water. The debris size of Ti-6Al-4V alloy in distilled water is greater than that in kerosene and compared with kerosene, the impulsive force of discharge in distilled water is smaller but more stable. Ekmekci et al. [31] presented an experimental work to measure residual stresses and hardness depth in EDM surfaces. Stresses are found to be increasing rapidly with respect to depth, attaining to its maximum value around the yield strength and then fall rapidly to compressive residual stresses in the core of the material since the stresses within plastically deformed layers are equilibrated with elastic stresses. In 2005, Sharma [32] investigated the potential of electrically conductive chemical vapor deposited diamond as an electrode for micro-electrical discharge machining in oil and water. While doing a comparative study on the surface integrity of plastic mold steel, Ekmekci [33] found that the amount of retained austenite phase and the intensity of micro cracks have found to be much less in the white layer of the samples machined in de-ionized water.

A new application in EDM power supply was designed to develop small size EDM systems by Casanueva. [34]. The proposed control achieves an optimum and stable operation using tap water as dielectric fluid to prevent the generation of undesired impulses and keep the distance between the electrode and the work piece within the optimum stable range. Studies conducted by Kang and Kim [35] in order to investigate the effects of EDM process conditions on the crack susceptibility of a nickel-based super alloy revealed that depending on the dielectric fluid and the post-EDM process such as solution heat treatment, cracks exist in recast layer could propagate into substrate when a 20% strain tensile force was applied at room temperature. When kerosene as dielectric, it was observed that carburization and sharp crack propagation along the grain boundary occurred after the heat treatment. However, using deionized water as dielectric the specimen after heat treatment underwent oxidation and showed no crack propagation behavior.

6.2 WATER WITH ADDITIVES

Koenig and Joerres [36] reported that a highly concentrated aqueous glycerin solution has an advantage as compared to hydrocarbon dielectrics when working with long pulse durations and high pulse duty factors and discharge currents, i.e. in the roughing range with high open-circuit voltages and positive polarity tool electrode. Leao and Pashby [37] found that some researchers have studied the feasibility of adding organic compound such as ethylene glycol, polyethylene glycol 200, polyethylene glycol 400, polyethylene glycol 600, dextrose and sucrose to improve the performance of deionized water. The surface of titanium has been modified after EDM using dielectric of urea solution in water [38]. The nitrogen element decomposed from the dielectric that contained urea, migrated to the work piece forming a TiN hard layer which

resulting in good wear resistance of the machined surface after EDM.

6.3 CONCLUSION

Water based EDM is more eco friendly as compared to conventional oil based dielectric. The material removal rate enhanced with use of water.

VII. MODELING

EDM process is influenced by many input factors. Various techniques viz. dimensional analysis, artificial neural network and thermal modeling are employed to predict the output of the process mainly the surface finish, tool wear and MRR.

7.1 DIMENSIONAL ANALYSIS

In 1979 Jeswani [39] used the dimensional analysis to predict the tool wear in EDM. The equation relates the volume of material eroded from the tool electrode to the energy of the pulse and density, thermal conductivity, specific heat and latent heat of vaporization of electrode material. Yahya and Manning [40] applied the method in analyzing MRR. The electrical and physical parameters related to the process are discharge pulse on time, gap voltage, sparking frequency, gap current and material Properties.

Results from the model show good agreement when compared to experimental finding. A semi empirical model of surface finish [41], MRR and tool wear [42] for various materials have been established by employing dimensional analysis. Using design of experiment method, the process parameter viz. peak current, pulse duration, electric polarity and material properties are identified. The final results show that the average error between the experiment and prediction was less than 10% for surface finish model and less than 20% for MRR. However, the relations between tool wear and discharge time under different electric polarity are seen to have inverse effect.

7.2 MATHEMATICAL MODEL

Dibitonto [43] presented a simple cathode erosion model for EDM. The Compton original energy balance for gas discharges is amended and the model uses the photoelectric effect as the dominant source of energy supplied to the cathode surface. Then, Patel [44] developed the anode erosion model which accepts power as boundary condition at anode interface and assumed to produce a Gaussian-distributed heat flux on the surface of anode material. The area upon which the flux is incident is assumed to grow with time. A model on variable mass and cylindrical plasma was introduced by Eubank [45]. It consists of three differential equation-one each from fluid dynamics, an energy balance and the radiation equation combined with plasma equations of state. Problems with the zero-time boundary conditions are overcome by an electron balance procedure. McGeough and Rasmussen [46] proposed a model for electro discharge texturing based on the effect of dielectric fluid and in particular the influence of change in the resistance in the dielectric during each voltage pulse. The theoretical predictions confirm the practical findings that the surface roughness in texturing is determined primarily by the peak current used and the length of the voltage 'on time'. Coguz [47] investigated on

surface profile of 2080 tool steel machined under varying machining parameters. It is found that surface roughness increases with increasing discharge current, pulse duration and dielectric flushing pressure. Surface profile information obtained is transferred to computer, digitized and then modeled in the form of Fourier series. Perez [48] presented a model for relative power dissipation by taking into account the different current emission mechanism and cathode space-charge characteristics valid for refractory and non-refractory materials. Tantra et al. [49] tested the validity of model proposed by Heuvelman for erosion strength of material to predict tool wear and their applicability to EDM process such as drilling of deep holes in turbine blades. The experimental results show the Heuvelman model does not show a direct correlation with the observation. Work proposed a combination of Taguchi method and Top[sis] to solve the multi-response parameter optimization problem in green electrical discharge machining. An analytical structure was developed to perform multi-criteria decision making. The responses were ranked based on the scores obtained by the summarization of final global preference weights. Triangular fuzzy numbers were used to assign preference values to the output responses. The optimum factor level combinations were identified based on the closeness coefficient values. The optimal machining performance for the green EDM was obtained for 4.5 A peak current (level 2), 261 μ m pulse duration (level 2), 40 mm dielectric level (level 1) and 0.5 kg/cm² flushing pressure (level 2). From analysis of the closeness coefficients, it was identified that the peak current was the most influential parameter in multi-performance characteristics used in this study. The computational and experimental effort needed to optimize these parameters was rather small. It was illustrated that the method was efficient and effective for multi-attribute decision making problems in green manufacturing. B. Izquierdo [50] presents an original thermal model capable of simulating discharge super position and representing EDM processed surfaces. The main conclusions obtained from this study are: Single-discharge modeling has been extensively studied in literature in order to characterize the EDM process. In this work, it has been shown that superposition of multiple discharges must be considered, since the amount of material removed per discharge increases (as much as 50%) as the operation progresses. This effect is related both to the stochastic nature of the process (discharge type and location) and the development of temperature fields on irregular surfaces. An original numerical model for simulation of the EDM process has been presented. The model generates EDM processed surfaces by calculating temperature fields inside the work piece using a finite difference-based approach, and taking into account the effect of successive discharges

7.3 ARTIFICIAL NEURAL NETWORK (ANN)

An attempt of modeling EDM process through ANN was carried out by Gopal and Rajurkar [51]. With machining depth, tool radius, orbital radius, radial step, offset depth, pulse on-time, pulse off-time and discharge current as the input parameters 9-9-2 size back propagation neural network was developed, experiments have been performed to check the validity of the ANN model and it can be concluded that the ANN model provides faster and

more accurate results. Tsai and Wang [52] have compared the ANN model on MRR. Six neural network models and neuro-fuzzy model for MRR have been established and analyzed. Results show that adaptive-network fuzzy inference system (ANFIS) is more accurate with an error 16.33%. In their further investigations, Tsai and Wang [53] have applied the same method to predict the surface finish. Results show that tangent sigmoid multi-layered perception (TANMLP), radial basis function network (RBFN),

Adaptive RBFN and ANFIS models have shown consistent results. Wang [54] combine the ANN and genetic algorithm to find an integrated solution to the problem of modeling and optimization of manufacturing processes. The error of the model is 5.6% for MRR and 4.98% for surface roughness. The modeling system established better knowledge about interaction between tool (graphite) and work piece (nickel alloy). An artificial feed forward neural network based on the Levenberg–Marquardt back propagation technique has been developed by Panda and Bhoi [50] to predict MRR. The model provides faster, more accurate results and performs well under the stochastic environment of actual machining conditions without understanding the complex physical phenomena exhibited in electro-discharge machining.

7.4 CONCLUSIONS

Various input parameters like current, pulse on, pulse off time; voltage has been taken for model formation to produce the proper combination for machining.

VIII. SUMMARY

The study reveals that a lot work has been done on estimation of tool wear and material removal rate in EDM. Researcher works on reduction of tool wear and enhancement of material removal rate by experimental investigation. use of powder mixed electrolyte, tool vibration, dry EDM use of gas instead of oil electrolyte, treatment on tool used for EDM, water based EDM and modeling techniques has been employed for increase of EDM efficiency. There is lot gap and no work has done on composite and ceramics materials. No CFD model for electrolyte flushing rate has been employed and there is lot work can be done on mathematical model of EDM.

REFERENCES

- [1] K.H. Ho, S.T. Newman, State of the art electrical discharge machining (EDM), *International Journal of Machine Tools & Manufacture* 43 (2003) 1287–1300.
- [2] C.J. Luis, I. Puertas, G. Villa, Material removal rate and electrode wear study on the EDM of silicon carbide, *Journal of Materials Processing Technology* 164–165 (2005) 889–896.
- [3] B. Bojorquez, R.T. Marloth, O.S. Es-Said, Formation of a crater in the work piece on an electrical discharge machine, *Engineering Failure Analysis* 9 (2002) 93–97.
- [4] J. Marafona, J.A.G. Chousal, A finite element model of EDM based on the Joule effect, *International Journal of Machine Tools & Manufacture* 46 (2005) 1–8.
- [5] H. Ramasawmy, L. Blunt, Effect of EDM process parameters on 3D surface topography, *Journal of Materials Processing Technology* 148 (2004) 155–164

- [6] F.-L. Zhao, Z.-Z. Lu, H. Wang, Z.-Q. Qian, Research on effecting mechanism of particles in powder-mixed EDM, Dalian Ligong Daxue Xuebao/Journal of Dalian University of Technology 45 (2005) 668–671.
- [7] P. Pecas, E. Henriques, Influence of silicon powder-mixed dielectric on conventional electrical discharge machining, International Journal of Machine Tools & Manufacture 43 (2003) 1465–1471.
- [8] Q. Yan Ming, L. You He, Powder-suspension dielectric fluid for EDM, Journal of Materials Processing Technology 52 (1995) 44–54.
- [9] Y.S. Wong, L.C. Lim, I. Rahuman, W.M. Tee, Near-mirror-finish phenomenon in EDM using powder-mixed dielectric, Journal of Materials Processing Technology 79 (1998) 30–40.
- [10] T. Yih-fong, C. Fu-chen, Investigation into some surface characteristics of electrical discharge machining SKD-11 using powder suspension dielectric oil, Materials Processing Technology 170 (2005) 385–391.
- [11] M.L. Jeswani, Effect of the addition of graphite powder to kerosene used as the dielectric fluid in electrical discharge machining, Wear 70 (1981) 133–139.
- [12] W.S. Zhao, Q.G. Meng, Z.L. Wang, The application of research on powder mixed EDM in rough machining, Journal of Material Processing Technology 129 (2002) 30–33.
- [13] Y.F. Tzeng, C.Y. Lee, Effects of powder characteristics on Electro discharge machining efficiency, International Journal of Advanced Manufacturing Technology 17 (2001) 586–592.
- [14] H.K. Kansal, S. Singh, P. Kumar, Application of Taguchi method for optimization of powder mixed electrical discharge machining, International Journal of Manufacturing Technology and Management 7 (2005) 329–341.
- [15] P. Pec “Electrical discharge machining using simple and powder-mixed dielectric: The effect of the electrode area in the surface roughness and topography.” journal of materials processing technology.vol.200, pp.250-258,2008.
- [16] Z.N. Guo, T.C. Lee, T.M. Yue, W.S. Lau, A study of ultrasonic aided wire electrical discharge machining, Journal of Materials Processing Technology 63 (1997) 823–828.
- [17] J.H. Zhang, T.C. Lee, W.S. Lau, X. Ai, Spark erosion with ultrasonic frequency, Journal of Materials Processing Technology 68 (1997) 83–88.
- [18] M. Ghoreishi, J. Atkinson, A comparative experimental study of machining characteristics in vibratory, Journal of Materials Processing Technology 120 (2002) 374–384.
- [19] Z.B. Yu, T. Jun, K. Masanori, Dry electrical discharge machining of cemented carbide, Journal of Materials Processing Technology 149 (2004) 353–357.
- [20] Darwin, Lal, D.M.,Nagarajan.“Optimization of cryogenic treatment of maximize the wear resistance of 18% Cr martensitic stainless steel.” Journal of. Material Processing. Technology, vol. 195,pp. 241-247, 2008.
- [21] Kalia, S. “Cryogenic processing: A study of materials of low temperatures.” journal of low temp physics,vol 158,pp. 934-945. 2010
- [22] Abdulkareem, S.Khan. “Reducing electrode wear ratio using cryogenic cooling during electrical discharge machining”. Int. J. Adv. Manuf. Technology, vol. 45, pp.1146-1151. 2009
- [23] Singh, R., Singh B., Singh..“ Experimental Investigation for Tool Life Enhancement using Cryogenic Treatment.” Journal of engineering science technology.vol. 4,pp. 1-4, 2010.
- [24] M. Dhananchezian. “Cryogenic Turning of AISI 304 Stainless Steel with Modified Tungsten Carbide Tool Inserts.”Elsevier cryogenic.vol. 51, pp.35-40,2011.
- [25] Q.H. Zhang, R. Du, J.H. Zhang, Q. Zhang, An investigation of ultrasonic-assisted electrical discharge machining in gas, International Journal of Machine Tools & Manufacture, DOI:10.1016/j.ijmachtools.2005.09.023
- [26] M.L. Jeswani, Electrical discharge machining in distilled water, Wear 72 (1981) 81–88.
- [27] S. Tariq Jilani, P.C. Pandey, Experimentnal investigations into the performance of water as dielectric in EDM, International Journal of Machine Tool Design and Research 24 (1984) 31–43.
- [28] W. Konig, F.-J. Siebers, Influence of the working medium on the removal process in EDM sinking, American Society of Mechanical Engineers, Production Engineering Division (Publication) PED 64 (1993) 649–658.
- [29] J.-P. Kruth, L. Stevens, L. Froyen, B. Lauwers, Study of the white layer of a surface machined by die-sinking electro-discharge machining, CIRP Annals—Manufacturing Technology 44 (1995) 169–172.
- [30] S.L. Chen, B.H. Yan, F.Y. Huang, Influence of kerosene and distilled water as dielectric on the electric discharge machining characteristics of Ti-6Al-4V, Journal of Materials Processing Technology 87 (1999) 107–111.
- [31] B. Ekmekci, O. Elkoca, A.E. Tekkaya, A. Erden, Residual stress state and hardness depth in electric discharge machining: de-ionized water as dielectric liquid, Machining Science and Technology 9 (2005) 39–61.
- [32] A. Sharma, M. Iwai, K. Suzuki, T. Uematsu, Potential of electrically conductive chemical vapor deposited diamond as an electrode for micro-electrical discharge machining in oil and water, New Diamond and Frontier Carbon Technology 15 (2005) 181–194.
- [33] B. Ekmekci, O. Elkoca, A. Erden, A comparative study on the surface integrity of plastic mold steel due to electric discharge machining, Metallurgical and Materials Transactions B: Process Metallurgy and Materials Processing Science 36 (2005) 117–124.
- [34] R. Casanueva, F.J. Azcondo, C. Brañ as, S. Bracho, Analysis design and experimental results of a high-frequency power supply for spark erosion, IEEE Transactions on Power Electronics 20 (2005) 361–369.
- [35] S.H. Kang, D.E. Kim, Effect of electrical discharge machining process on crack susceptibility of nickel based heat resistant alloy, Materials Science and Technology 21 (2005) 817–823.

- [36] W. Koenig, L. Joerres, Aqueous solutions of organic compounds as dielectric for EDM sinking, *CIRP Annals Manufacturing Technology* 36 (1987) 105–109.
- [37] F.N. Leao, I.R. Pashby, A review on the use of environmentally friendly dielectric fluids in electrical discharge machining, *Journal of Materials Processing Technology* 149 (2004) 341–346.
- [38] B.H. Yan, H.C. Tsai, F.Y. Huang, The effect in EDM of a dielectric of a urea solution in water on modifying the surface of titanium, *International Journal of Machine Tools & Manufacture* 45 (2005) 194–200.
- [39] M.L. Jeswani, Dimensional analysis of tool wear in electrical discharge machining, *Wear* 55 (1979) 153–161.
- [40] A. Yahya, C.D. Manning, Determination of material removal rate of an electro-discharge machine using dimensional analysis, *Journal of Physics D: Applied Physics* 37 (2004) 1467–1471.
- [41] K.-M. Tsai, P.-J. Wang, Semi-empirical model of surface finish on electrical discharge machining, *International Journal of Machine Tools & Manufacture* 41 (2001) 1455–1477.
- [42] P.-J. Wang, K.-M. Tsai, Semi-empirical model on work removal and tool wear in electrical discharge machining, *Journal of Materials Processing Technology* 114 (2001) 1–17.
- [43] D.D. DiBitonto, P.T. Eubank, M.R. Patel, M.A. Barrufet, Theoretical models of the electrical discharge machining process. I. A simple cathode erosion model, *Journal of Applied Physics* 66 (1989) 4095–4103.
- [44] M.R. Patel, M.A. Barrufet, P.T. Eubank, D.D. DiBitonto, Theoretical models of the electrical discharge machining process. II. The anode erosion model, *Journal of Applied Physics* 66 (1989) 4104–4111.
- [45] P.T. Eubank, M.R. Patel, M.A. Barrufet, B. Bozkurt, Theoretical models of the electrical discharge machining process. III. The variable mass, cylindrical plasma model, *Journal of Applied Physics* 73 (1993) 7900–7909.
- [46] J.A. McGeough, H. Rasmussen, A theoretical model of electro discharge texturing, *Journal of Materials Processing Technology* 68 (1997) 172–178.
- [47] C. C-og̃ un, B. Kocabas, A. Õ zgedik, Experimental and theoretical investigation of workpiece surface profiles in electrical discharge machining (EDM), *Journal of the Faculty of Engineering and Architecture of Gazi University* 19 (2004) 97–106.
- [48] R. Perez, H. Rojas, G. Walder, R. Flukiger, Theoretical modeling of energy balance in electro erosion, *Journal of Materials Processing Technology* 149 (2004) 198–203.
- [49] N.Y. Tantra, F. Leao, I.R. Pashby, Evaluating theoretical equations to predict wear in electro-discharge machining, *Proceedings of the First International Conference and Seventh AUN/SEED-Net Field wise Seminar on Manufacturing and Material Processing*, 14–15 March 2006, Kuala Lumpur, 151-155 (ISBN:983-42876-0-7).
- [50] B. Izquierdo. “A numerical model of the EDM process considering the effect of multiple discharges.” *International Journal of Machine Tools & Manufacture*.vol. 49 PP 220–229(2009)
- [51] I. Gopal, K.P. Rajurkar, Artificial Neural Network approach in modeling of EDM process, *Intelligent Engineering Systems Through Artificial Neural Networks* 2 (1992) 845–850.
- [52] K.-M. Tsai, P.-J. Wang, Comparisons of neural network models on material removal rate in electrical discharge machining, *Journal of Materials Processing Technology* 17 (2001) 111–124.
- [53] K.-M. Tsai, P.-J. Wang, Predictions on surface finish in electrical discharge machining based upon neural network models, *International Journal of Machine Tools & Manufacture* 41 (2001) 1385–1403.
- [54] K. Wang, H.L. Gelgele, Y. Wang, Q. Yuan, M. Fang, A hybrid intelligent method for modeling the EDM process, *International Journal of Machine Tools & Manufacture* 43 (2003) 995–999.
- [55] D.K. Panda, R.K. Bhoi, Artificial neural network prediction of material removal rate in electro discharge machining, *Materials and Manufacturing Processes* 20 (2005) 645–672

Top-K Retrieval Cbir Systems

Pradeep yadav Mtech (S.E)

Poornima college of engineering,jaipur

ABSTRACT: Multimedia data is now one of the widely used information types in all the fields as the fast development of computer techniques has been made traditional database systems based on text information have the limitations when applied to multimedia information.

One of the main problems the researchers highlighted was the difficulty of locating a desired image in a large and varied collection. While it is perfectly feasible to identify a desired image from a small collection simply by browsing, more effective techniques are needed with collections containing thousands of items.

In this paper, the problem involves entering an image as a query into a software application that is designed to employ CBIR techniques in extracting visual properties, and matching them. This is done to retrieve images in the database that are visually similar to the query image.

I. INTRODUCTION

The purpose of this paper is to provide an overview of the functionality of temporary image retrieval systems in terms of technical aspects: querying, relevance feedback, features, matching measures, indexing data structures, and result presentation. It compares specific systems, rather than general architectures, and provides a basis for (or defense against) statements like “this or that system already does what your system does”. It also is a thorough foundation for claims that most systems use low Level features, and few use high level semantic meaningful features.

Images now play a crucial role in fields as diverse as medicine, journalism, advertising, design, education and entertainment. Technology, in the form of inventions such as photography and television, has played a major role in facilitating the capture and communication of image data. But the real engine of the imaging revolution has been the computer, bringing with it a range of techniques for digital image capture, processing, storage and transmission which would surely have startled even pioneers like John Logie Baird. The involvement of computers in imaging can be dated back to 1965, which demonstrated the feasibility of computerized creation, manipulation and storage of images, though the high cost of hardware limited their use until the

mid-1980s. Once computerized imaging became affordable, it soon penetrated into areas traditionally depending heavily on images for communication, such as engineering, architecture and medicine. Photograph libraries, art galleries and museums, too, began to see the advantages of making their collections available in electronic form. The creation of the World-Wide Web in the early 1990s, enabling users to access data in a variety of media from anywhere on the planet, has provided a further massive stimulus to the exploitation of digital images. The number of images available on the Web was recently estimated to be between 10 and 30 million [3].

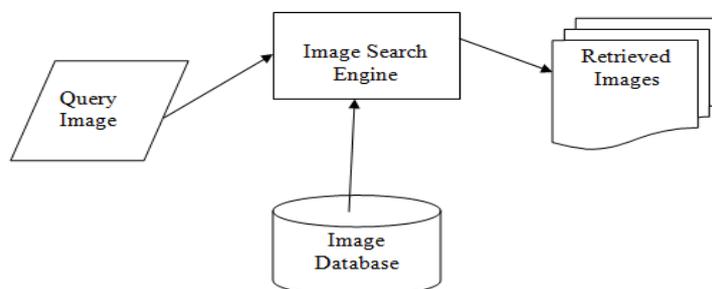
A. Digital Image Representation

The term monochrome image or simply image refers to a two dimensional light intensity function $f(i, j)$, where 'i' and 'j' denote the spatial coordinates and the value of 'f' at any point (x, y) is proportional to the brightness level (gray level) of the image at that point. A digital image is an image that has been quantized both in spatial coordinates and brightness. A digital image can be considered a matrix whose row and column indices identifying a point in the image and the corresponding matrix element value identify the gray level at that point. The elements of such a digital array are called image elements, picture elements, pixels or pels.

B. Need for Image Retrieval: One of the main problems the researchers highlighted was the difficulty of locating a desired image in a large and varied collection. While it is perfectly feasible to identify a desired image from a small collection simply by browsing, more effective techniques are needed with collections containing thousands of items. Journalists requesting photographs of a particular type of event, designers looking for materials with a particular colour or texture, and engineers looking for drawings of a particular type of part, all need some form of access by image content.

The basic concept of image retrieval is to find out whether an image or the image-database contains the query pattern given by the user on the basis of some similarity function. The basic block diagram of Image Retrieval is shown

below in figure.1



II. RELATED WORK

CONTENT-BASED image retrieval (CBIR) has received much attention in the last decade, which is motivated by the need to efficiently handle the immensely growing amount of multimedia data. In a typical CBIR system, low-level visual image features (e.g., color, texture, and shape) are automatically extracted for image descriptions and indexing purposes. To search for desirable images, a user presents an image as an example of similarity, and the system returns a set of similar images based on the extracted features. In CBIR systems with relevance feedback (RF), a user can mark returned images as positive or negative, which are then fed back into the systems as a new, refined query for the next round of retrieval. The process is repeated until the user is satisfied with the query result. Such systems are effective for many practical CBIR applications. There are two general types of image search:

Category search

The goal of target search is to find a specific (target) image, such as a registered logo, a historical photograph, or a particular painting. The goal of category search is to retrieve a given semantic class or genre of images, such as scenery images or skyscrapers. In other words, a user uses target search to find a known image. In contrast, category search is used to find relevant images the user might not be aware ahead of time. We focus on category search in the next subsequent chapters of this project work.

A. Indexing Techniques

Hierarchical clustering

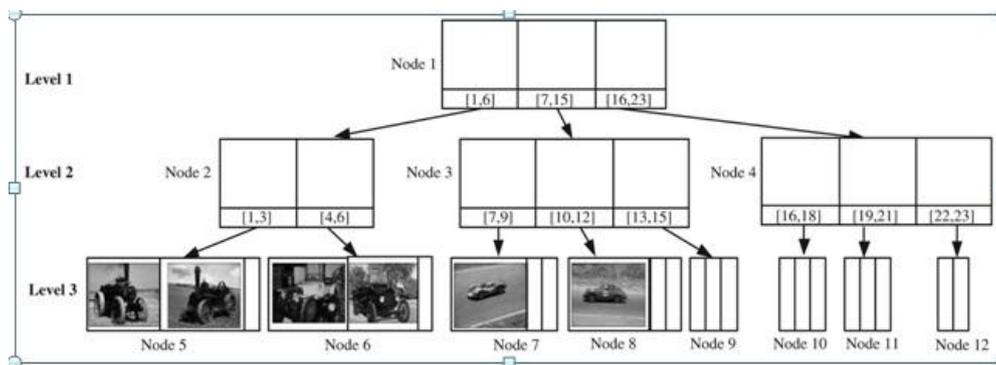


Figure. 2.1. The Hierarchical clustering index structure.

A hierarchical clustering technique, similar to the R*-tree [4], is used to organize the entire image database into a hierarchical tree structure. With each node in this hierarchy representing a cluster, the original node structure of the R*-tree is extended to include also information to identify the images in their children nodes. The hierarchical clustering is constructed as follows: When a new element (i.e., an image represented as a high-dimensional point) is inserted into the tree, this element is placed in a leaf node that requires the least enlargement in its bounding box, and a leaf node's MBB is based on all dimensions of its contained image points. If a leaf node overflows, this node is split (i.e., a portion of its entries is removed from the node

and reinserted into the tree), and such splits propagate up the tree[4].

The R* Tree

Two orthogonal issues in CBIR research are efficiency and accuracy. For instance, indexing techniques, such as R*-tree[4], may improve the efficiency of the search process. Their retrieval accuracy, however, depends on the effectiveness of the visual features used to characterize the database images. An effective CBIR system, therefore, needs to have both an efficient search mechanism and an accurate set of visual features.

The R-tree, one of the most popular access methods for rectangles, is based on the heuristic optimization of the area of the enclosing rectangle in each inner node. By running numerous experiments in a standardized test bed under highly varying data, queries and operations, we were able to design the R*-tree[4] which incorporates a combined optimization of area, margin and overlap of each enclosing rectangle in the directory.

- From a practical point of view the R*-tree is very attractive because of the following two reasons
- It efficiently supports point and spatial data at the same time.
 - Its implementation cost is only slightly higher than that of other R-trees.

An important requirement for a spatial access method is to handle both spatial objects and point objects efficiently. Points can be considered as degenerated rectangles and in most applications rectangles are very small relatively to the data space.

We traverse the tree in a postorder fashion. In the original R*-tree, an internal node contains an array of node entries. Each node entry is a pair (mbb, node-id), where mbb is the MBB that spatially contains the MBBs in the child node, with node-id as the child node address. In our index structure, each node entry is extended to be a tuple (mbb, node-id, imageIDrange), where imageID-range refers to the range of image identifications contained in the pointed child node and imageID-range C [1,|S|]. Let us describe how to augment the structure illustrated in Figure 2.1. We start from the root node (i.e., Node 1) which has three node entries. We first visit the first node entry which

points to Node 2. Node 2 has two node entries, pointing to leaf nodes 5 and 6 in order. Our depth-first traversal leads us to Node 5, which contains three image points. Then, we set the imageIDrange of the first entry in Node 2 to [1, 3], and each image contained in this node can randomly pick an exclusive ID within this range. That is, the three images in Node 5 can be assigned IDs 1, 2, and 3, respectively.

iterations) when they again fall within the search range of the current iteration. This strategy leads to the following disadvantages:

Local maximum trap problem

In this type problem there is no guarantee that the target can be found. The search operation generally takes several iterations of RF to examine a number of regions in the feature space, before it reaches the target image. During this iterative process, the search advancement might get trapped in a region as illustrated in Figure 2.2. It shows s

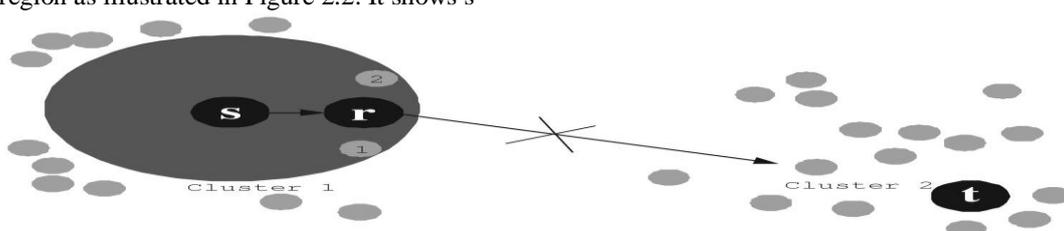


Figure. 2.2. Local maximum trap in existing approaches.

scenario, the search process is trapped in this local region, and can never reach the target point p_t . Although, the system can escape the local maximum trap with a larger k , it is difficult to guess a proper threshold ($k = 14$ in this example). Consequently, the user might not even know a local maximum trap is occurring.

Slow convergence problem

Including previously examined images in the computation of the current centroid results in repeat retrieval of some of the images. This prevents a more aggressive movement of the search in the feature space. This drawback is illustrated in Figure. 2.3, where $k = 3$. It shows that it takes six iterations for the search operation to reach the target point p_t . This slow convergence incurs

longer search time, and significant computation and disk access overhead.

III. Proposed System Architecture

In most approaches image features are extracted, stored in a database and compared with the features of a particular search image.

In CBIR, each image in the database has its features extracted and compared to the features of the query image. It involves two steps:

Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.

Matching: The second step involves matching these features to yield a result that is visually similar.

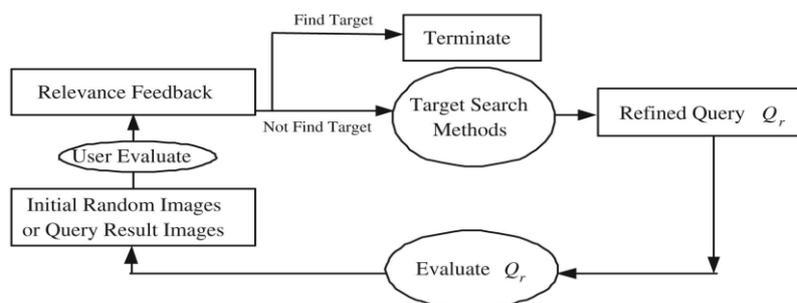


Figure. 3.1. Overview of the target search systems

In the architecture shown in figure.3.1 the user first submits a query image to the system. The features of the query image is calculated. Then the system search the feature database with the help of the target search methods described in the previous chapter. In each iteration the system extracts some refined images based on the relevance

feedback. If the target image is found from the result images, then the process terminates otherwise the user again selects some of the images from the result image set relative to the target image. Then the Query is prepared and the search process is continued in an iterative fashion till the target image is found.

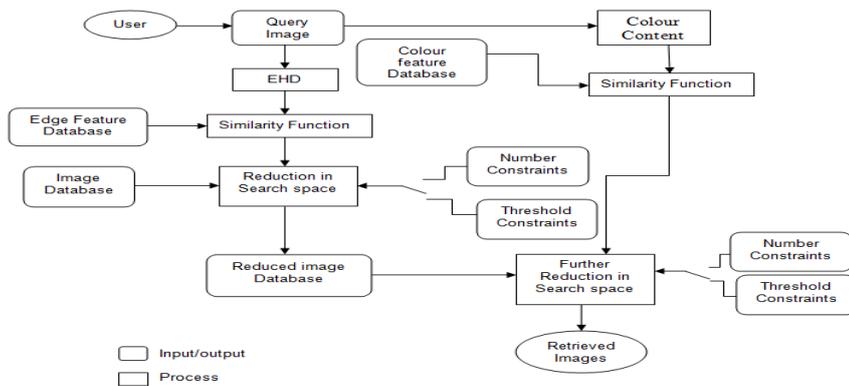


Figure 3.4: Hierarchical Architecture

This architecture is used to retrieve similar images based on edge features and colour features of the query image. Unlike combining the features, a hierarchical retrieval system is proposed which improves the retrieval performance and minimizes the search time.

When the user submits a query image the system asks the user to enter the desired colour (foreground colour). Hence the background colour is automatically neglected. Edge histogram containing 5 types of edges (point, horizontal, vertical, 45° diagonal, 135° diagonal) is extracted and compared with the edge feature database. Relevant images are retrieved in the order of similarity.

The number of images retrieved by the system depends on the threshold or number constraint given to the system.

IV. Features Used For Proposed Hierarchical Method

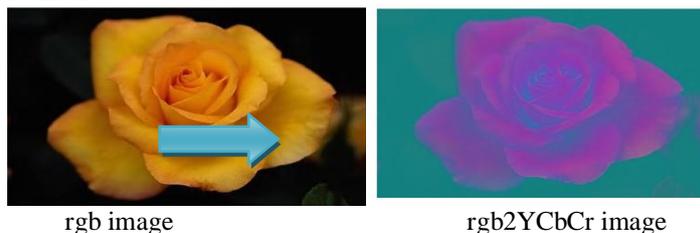
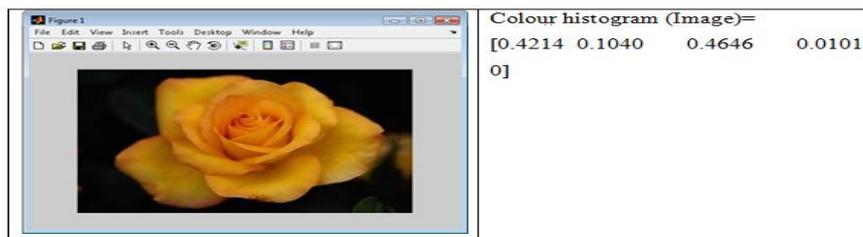


Figure 4.6: rgb2YCbCr conversion of an image

The colour histograms for the images are computed by quantizing the colours within the image. Each colour image is split up in to 3 two-dimensional images such as Luminance (Y), Blue chrominance (Cb) and Red Chrominance (Cr) respectively. Number of bins in the colour histogram can be set to obtain the feature vector of desired size [5].



4.4.2. Colour Content Histogram:

B. Colour Content Histogram

Colour content shows the dominant colour present in the image. The entire RGB colour space is described
 Table 4.2: Color Content look-up Table

using small set of colour content categories. This is summarized into a colour content look up table.

Color Content	R	G	B

Black	<50	<50	<50
Blue	<50	<100	>100
Red	>150	<50	<100
Yellow	>100	>100	<50
Magenta(Violet)	>80	<50	>80
Cyan(Indigo)	<50	>150	>100
Green	<100	>80	<50
Gray	<50 & >150	<50& >150	<50 & >150
White	>180	>180	>180

Retrieval has been classified as accurate if for a given query image the system retrieves the perceptually (to human) most similar image as the topmost retrieval. For each query, images similar to the query image are manually listed from the database by their dominant colours present in the query. To evaluate the retrieval performance, the standard evaluation method, i.e., precision-recall pair [49] is used. Precision (P) is defined as the ratio of number of relevant images retrieved (N) to the total number of images retrieved (K). Recall (R) is defined as the ratio of number of relevant images retrieved (N) to the total number of relevant images in the whole database (T).

$$\text{Precision } P = \frac{N}{K} \dots\dots\dots (6.1)$$

$$\text{Recall } R = \frac{N}{T} \dots\dots\dots (6.2)$$

Where
 N is the number of relevant images retrieved
 K is the total no. of images retrieved.
 T is the total no. of relevant images in the database.

V. EXPERIMENTS AND RESULTS

The image database used in this experiment is downloaded from WBIIS database [48]. The database consists of 1000 images. The image set comprises 100 images in each of 10 categories. The images are of the size 256x384 or 384x256. The types of images include flowers, buses, animals, constructions, tribal photos, natural sceneries, beaches, fruits etc. Accuracy of the retrieval scheme is calculated by Precision & recall graph.

image.orig\643.jpg image.orig\602.jpg image.orig\634.jpg image.orig\624.jpg image.orig\668.jpg



image.orig\618.jpg image.orig\608.jpg image.orig\650.jpg image.orig\672.jpg image.orig\677.jpg



image.orig\696.jpg image.orig\61.jpg image.orig\635.jpg image.orig\669.jpg image.orig\648.jpg



image.orig\693.jpg image.orig\657.jpg image.orig\601.jpg image.orig\680.jpg image.orig\663.jpg



Figure 6.1(c): Retrieved Images

Table-6.1: No. of retrieved images vs. No. of relevant retrieved images

Method	K=1	K=8	K= 20
Euclidean Distance	N=1	N=3	N=10
Normalized Absolute Difference	N=1	N=5	N=10

Min-Max Ratio	N=1	N=6	N=12
---------------	-----	-----	------

From the above table the precision recall Graph is plotted below.

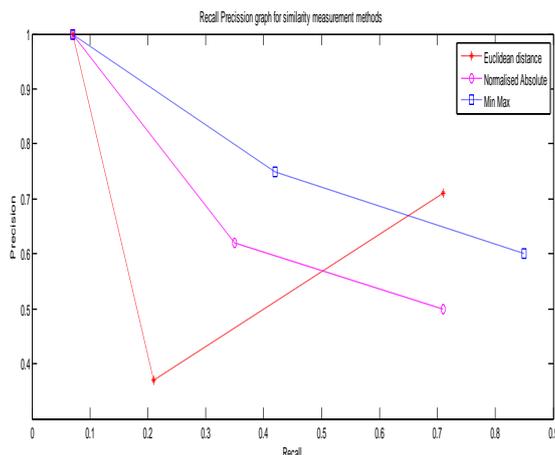


Figure-6.2: Precision and Recall Graph of the different types of similarity measurement method

From the above observation it can be concluded that min-max ratio is better than the other similarity measurement method. For further experiments the Min-Max ratio is taken for the similarity measurement method.

VI. CONCLUSION

In this paper, we proposed a new paradigm inspired in one of the ways people can find objects. To implement and test this approach, two level descriptors are selected and organized into a hierarchical structure. The proposed hierarchical retrieval method is less time consuming than the combined features method. In the proposed hierarchical method colour feature and edge features are considered in different stages. For colour features we have considered Colour Content Histogram (CCH). The proposed technique of dominant colour identification based on CCH is a meaningful technique to retrieve the images based on colour. Selection of prominent colour by the use of, helps in segmenting the foreground from the background of the image. Hence the colour content feature is giving better retrieval results than the colour histogram.

REFERENCES

- [1]. D. Liu, K.A. Hua, K. Vu, and N. Yu, "Fast Query Point Movement Techniques for Large CBIR Systems IEEE Transactions on Knowledge and data Engineering, VOL. 21, NO. 5, MAY 2009.
- [2]. P. R. Snoeren and N. Karssemeijer. "Thickness correction of mammographic images by means of a global parameter model of the compressed breast.", IEEE Transactions on Medical Imaging, 23(7): pp 799-806, 2004.
- [3]. Sclaroff S et al (1997) "ImageRover: a content-based image browser for the World-Wide Web" in Proceedings of IEEE Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico, June 1997, pp 2-9.
- [4]. Norbert Beckmann, Hans-Peter Kriegel, Ralf Schneider, Bernhard Seeger "The R*-tree: An Efficient and Robust Access Method for Points and Rectangles" ACM 1990.
- [5]. Eakins, J P (1996) "Automatic image content retrieval – are we getting anywhere?" Proceedings of Third International Conference on Electronic Library and Visual Information Research (ELVIRA3), De Montfort University, Milton Keynes, pp 123-135
- [6]. Eakins J P (1998) "Techniques for image retrieval" Library and Information Briefings, in press
- [7]. Stian Edvardsen, "Classification of Images using Colour, CBIR Distance Measures and Genetic Programming. An evolutionary Experiment." (2006) Norwegian University of Science and Technology.
- [8]. Dengsheng Zhang, "Improving Image retrieval Performance by using both Colour and Texture Features.", Proceedings of the third International Conference on Image and Graphics (ICIG'04). (2004) IEEE.
- [9]. W.Niblack et al. "The QBIC project: Querying Images By Content using Colour, Texture and Shape.", SPIE Conf. on storage and Retrieval for image and video Databases, vol.1908, San Jose, CA, pp.173-187, 1993.
- [10]. M. Swain and D. Ballard, "colour Indexing", International Journal of Computer Vision, 7(1): pp 11-32, 1991
- [11]. Kato T (1992) "Database architecture for content-based image retrieval" in Image Storage and Retrieval Systems, Proc SPIE 1662, pp112-123.
- [12]. Gudivada V N and Raghavan V. V. , "Content-based image retrieval systems" IEEE Computer 28(9), pp18-22.
- [13]. Jain, R (1995) "World-wide maze" IEEE Multimedia 2(3), pp 3
- [14]. Enser P G B (1995) "Pictorial information retrieval" Journal of Documentation, 51(2), pp126-170

- [15]. Flickner, M et al "Query by image and video content: the QBIC system" IEEE Computer 28(9), pp23-32
- [16]. Gupta, A et al "The Virage image search engine: an open framework for image management" in Storage and Retrieval for Image and Video Databases IV, Proc SPIE 2670, pp 76-87
- [17]. P. J. Werkhoven, J. M. Schraagen, and P. A. J. Punte. "Seeing is believing: Communication performance under isotropic teleconferencing conditions. Displays", 22(4): pp 137-149, 2001.
- [18]. Remco C. Veltkamp, Mirela Tanase, "Content-Based Image Retrieval Systems: A Survey", Technical Report UU-CS-2000-34, October 2000.
- [19]. T. Kato. "Database architecture for content-based image retrieval.", Proceedings of SPIE Image Storage and Retrieval Systems, volume 1662, pages 112-123, San Jose, CA, USA, February 1992.
- [20]. Del Bimbo, A., "Visual Information Retrieval". Morgan Kaufmann Publishers, Inc
- [21]. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. "Content-based image retrieval at the end of the early years." IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12): pp1349-1380, 2000.
- [22]. Y. Rui, T. S. Huang, and S.-F. Chang. "Image retrieval: Past, present, and future." Journal of Visual Communication and Image Representation, 10:1.23, 1999.
- [23]. L. Schomaker, L. Vuurpijl, and E. de Leau. "New use for the pen: outline-based image queries." In Proceedings of the 5th IEEE International Conference on Document Analysis, pages 293-296, Piscataway (NJ), USA, 1999.
- [24]. L. H. Armitage and P. G. B. Enser. "Analysis of user need in image archives." Journal of Information Science, 23(4):pp 287-299, 1997.
- [25]. G. Ciocca and R. Schettini. "Using a relevance feedback mechanism to improve content based image retrieval." In D. P. Huijsmans and A. W. M. Smeulders, editors, Visual Information and Information Systems: Third International Conference (VISUAL'99), volume 1614 of Lecture Notes in Computer Science, pages 107-114. Springer-Verlag GmbH, 1999.
- [26]. N. Ikonomakis, K. N. Plataniotis, and A. N. Venetsanopoulos. "User interaction in region-based color image segmentation". In D. P. Huijsmans and A. W. M. Smeulders, editors, Visual Information and Information Systems: Third International Conference (VISUAL'99), volume 1614 of Lecture Notes in Computer Science, pages 99-106. Springer- Verlag GmbH, 1999.
- [27]. C. J Jørgensen. Access to pictorial material: "A review of current research and future prospects". Computers and the Humanities, 33(4):pp-293-318., 1999.

Optical Absorption and Fluorescence Spectral Analysis of Nd³⁺ Doped Bismuth Boro-Silicate Glasses

Sunil Bhardwaj^a, Rajni Shukla^a, Sujata Sanghi^b, Ashish Agarwal^b, Inder Pal^b

^aDepartment of Physics, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonapat-131093, Haryana, India

^bDepartment of Applied Physics, Guru Jambheshwar University of Science and Technology, Hisar-125001, Haryana, India

Abstract: Glasses having compositions $20\text{B}_2\text{O}_3.(79.5-x)\text{Bi}_2\text{O}_3.x\text{SiO}_2$ ($10 \leq x \leq 40$, mol%) doped with 0.5 mol% of Nd³⁺ ions were prepared by melt quench technique. The amorphous nature of the glasses was confirmed by X-ray diffraction studies. The spectroscopic properties of Nd³⁺ ions in borate bismuth silicate glasses as a function of bismuth oxide were investigated using optical absorption and fluorescence spectra. The Judd-Ofelt theory has been employed to calculate transitions probability from the data of absorption cross-section of several f-f transitions. The intensity parameters Ω_2 , Ω_4 , and Ω_6 are determined by applying least square analysis method. The variation of Ω_2 and Ω_6 with Bi₂O₃ content has been attributed to changes in the asymmetry of the ligand field at the rare earth ion site and to the changes in the rare earth oxygen (RE-O) covalency. The variation of Ω_4 with Bi₂O₃ content has been attributed to rigidity of the samples. Using the Judd Ofelt intensity parameters the other radiative properties like radiative transition probability, radiative life time, branching ratio and the stimulated emission cross-sections of prepared BBSN glasses have been calculated. The main fluorescence transition ${}^4\text{F}_{3/2} \rightarrow {}^4\text{I}_{11/2}$ of Nd³⁺ ion has been investigated as a function of Bi₂O₃ in host glass.

Keywords: Fluorescence properties, Heavy metal oxides, Judd-Ofelt theory, Optical properties, Rare earth ions

I. INTRODUCTION

In recent decades an increasing interest in rare earth doped glasses from the viewpoint of their spectroscopic properties and technological applications in many fields is observed. In fact, they are suitable materials for photonic devices such as planar optical waveguides, optical fiber amplifiers, wave guide laser etc [1-3]. Ever since the first solid state laser was demonstrated in 1961 by Snitzer in Nd³⁺ doped glasses, a lot of efforts are being done for the development of laser glasses. Neodymium doped glasses have attracted attention as they act as key element for optical amplifier around 1.3 μm and for higher power laser applications around 1.05 μm [4]. The outer shell configuration of neodymium is $4f^{12} 5s^2 5p^6 6s^2$. The transition probability between 4f states, however are sensitive to the ions surrounding the rare earth ions.

In general, the optical and spectroscopic properties of rare earth ions are strongly dependent on host materials. The host glass materials should have high refractive index with good chemical and thermal stability along with low

melting temperature of heavy metals in order to become more practically useful industries. Many potential host materials for rare earth ions have been developed [5]. One of the preferred host materials is oxide glasses which are chemically durable, thermally stable, optically transparent at the excitation and lasing wavelengths [6]. Glasses based on heavy metal oxides (such as Bi₂O₃) have received increased attention due to their manifold possible applications in the field of glass ceramics, layers for optical, optoelectronics devices, thermal and mechanical sensors, reflecting windows etc [7]. The large polarizability of bismuth makes it suitable for possible non linear optical devices and environmental guide lines [8]. Wide transmitting window in the optical region having sharp cut-off in both UV-VIS and IR region make these glasses useful in spectral devices [9].

Further, Judd-Ofelt theory has been employed to calculate the intensities of f-f transition in the absorption spectra of the rare earth ion. Experimental investigations of absorption spectra of rare earth ions in different matrices have revealed that the oscillator strengths for the specific transitions of rare earth ions depend more strongly on the structural features of materials and on the chemical nature of the ligands surrounding the rare earth ion as compared to other transitions. Once the experimental values of the oscillator strengths are determined from the integrated absorption coefficient, the intensity parameters can be obtained. When the phenomenological parameters Ω_λ are calculated, it is possible to derive the strength of any absorption or emission transition, as well as the stimulated emission cross-section and fluorescence branching ratio. Such spectral studies give valuable information about the structure and bonding in glass and radiative and non radiative properties of rare earth ions doped in glass matrix. This information is essentially required while developing new optical devices [10].

II. Theoretical details

Judd-Ofelt [11-12] theory has been used to investigate radiative nature of trivalent rare earth ions in a variety of laser host materials. The intensity parameter, radiative life time, and branching ratio are calculated with refractive index using Judd-Ofelt analysis. The experimental oscillator strength was calculated from the absorption spectra by using the equation.

$$f_{\text{meas.}} = 4.32 \times 10^{-9} \int \epsilon(\nu) d\nu \quad (1)$$

where $\epsilon(\nu)$ is the molar absorptivity (cm^2) of absorption band at energy $\nu(\text{cm}^{-1})$. The experimental oscillator strengths were used to obtained values of the intensity parameters ($\lambda = 2, 4, 6$) following the standard least square fitting method. The validity of the fitting is examined by comparing the experimental and theoretical oscillator strengths. The Judd-Ofelt intensity parameters (Ω_λ) are phenomenological characteristics for the influence of surrounding environment of the rare earth ions as they contain implicitly the crystal field terms and the radial wave function. According to the Judd-Ofelt theory the oscillator strength for an electronic transition from the initial manifold (SL, J) to the final manifold (S'L', J') is given by the expression.

$$f_{\text{cal}} = \frac{8\pi^2 mc}{3h(2j+1)\lambda_p} \left(\frac{(n^2+2)^2 S_{ed}}{9n} \right) \quad (2)$$

Where c represents the velocity of light, n represent refractive index, λ_p represent the absorption peak wavelength, e and m the electron charge and mass, S_{ed} represents the line strengths for the induced electric dipole transition, which is given by:

$$S_{ed} = \sum_{\lambda=2,4,6} \Omega_\lambda \left| \langle SLJ \| U^\lambda \| S'L'J' \rangle \right|^2 \quad (3)$$

The spontaneous emission probability for the radiative decay, A_{rad} (sec^{-1}), from the initial manifold (SLJ) to the final manifold (S'L'J') is given by the expression.

$$A_{rad} = \frac{64\pi^4 e^2 \bar{\nu}^3 n(n^2+2)^2}{3h(2j+1)9} \times S_{ed} \quad (4)$$

Where $U^{(\lambda)}$ ($\lambda=2, 4, 6$) are the matrix elements of unit tensor operators and had been calculated by Carnall et al. The total probability, A_T , obtained by carrying out the summation of all the transitions to the final state bj' is given by

$$A_T = \sum A_{rad} \quad (5)$$

The fluorescence branching ratio is given by

$$\beta_r = \frac{A_{rad}}{A_T} \quad (6)$$

The radiative life time is represented by

$$\tau_r = \frac{1}{A_T} \quad (7)$$

The induced emission cross-section for each transition could be estimated from the emission spectra by [9].

$$\sigma = \frac{\lambda_p^4 A_{rad}}{8\pi c n^2 \Delta\lambda} \quad (8)$$

Where λ_p is peak wavelength and $\Delta\lambda$ is the full width at half maxima of the fluorescent peak for different transitions obtained from emission spectra.

III. Experimental

Glasses having compositions $20\text{B}_2\text{O}_3.(79.5-x)\text{Bi}_2\text{O}_3.x\text{SiO}_2$ ($10 \leq x \leq 40$, mol%) doped with 0.5 mol% of Nd^{3+} ions were prepared by melt quench technique. Appropriate amount of AR grade chemicals (B_2O_3 , Bi_2O_3 , SiO_2 and Nd_2O_3) having purity above 99.99% were weighed on 0.001% accuracy and mixed thoroughly. The raw mixed materials were melted in a muffle furnace in air (at 1150°C for 40 min). The crucible was shaken frequently after every 10 minutes for the homogeneous mixing of all the constituents. The melt was quenched at room temperature by pouring between two stainless steel plates. The quenched samples were annealed to minimize the internal strain and then allowed to cool slowly to room temperature. The exact composition along with the codes of the samples is given in Table 1. The samples were polished for spectral and other investigations. The amorphous nature of the glasses was confirmed by recording the X-ray diffraction pattern using Mini Flex Desktop X-ray Diffractometer with $\text{Cu-K}\alpha$ line of wavelength at the scanning rate of $2^\circ/\text{min}$ and 2θ was varied from 10° to 80° .

The refractive index (n) of the prepared samples was measured by the Brewster angle method using He-Ne laser (633 nm). The density (D_g) of all the glasses were measured by using Archimedes principle with xylene as immersing liquid. The relation used is

$$D_g (\text{gm/cm}^3) = \frac{W_a}{W_a - W_b} D_x \quad (9)$$

Where W_a is the weight of glass sample in air, W_b is the weight of glass sample when immersed in xylene and D_x is the density of xylene (0.86 gm/cm^3).

The Optical absorption spectra of all the polished samples were recorded on a Varian-Carry 5000 spectrophotometer in the range 300-3200 nm. The emission spectra were recorded using Fluoro-max-3 Fluorimeter with Xe arc lamp as the excitation source at wavelength of 800 nm.

IV. RESULT AND DISCUSSION

4.1 Physical parameters

Various physical properties of all glasses as a function of Bi_2O_3 content were determined from the experimental data and are listed in Table 1. It is clearly observed that density decrease monotonously when Bi_2O_3 is replaced by SiO_2 in the glass. This may be due to greater molar mass of Bi_2O_3 (438.96) than that of SiO_2 (60.08). The decrease in molar volume on the other hand shows that addition of SiO_2 may contract the structure of the loose network in the studied glass. Also the smaller values of radii and bond length of SiO_2 than that of Bi_2O_3 results in decrease in the volume of these glasses

During recent years the optical basicity has been successfully used to correlate a range of properties of glass with chemical composition. The optical basicity of a material can be experimentally determined using UV probe ion spectroscopy. It was found that the frequency of the UV absorption band shifts gradually with composition. The

optical basicity (Λ) addresses the ability of oxide glass in contributing the negative charges in the glass matrix. In other words it defines the electron donating power of the oxygen in the oxides glass. The theoretical optical basicity can be calculated by the equation proposed by Duffy and Ingram

$$\Lambda_{th} = X_1\Lambda_1 + X_2\Lambda_2 + X_3\Lambda_3 + \dots + X_n\Lambda_n \quad (10)$$

Where X_1, X_2, \dots, X_n are equivalent fraction based on the amount of oxygen each oxide contributes to the overall glass stoichiometry and $\Lambda_1, \Lambda_2, \Lambda_3, \dots, \Lambda_n$ are basicities assigned to the individual oxides.

The optical basicity of Bi_2O_3 is higher than that of SiO_2 therefore basicity decreases with decrease of Bi_2O_3 in the host glass.

Electronic polarizability, α , has important chemical implications and has been the subject of several studies. It can be obtained experimentally from refractivity measurements using the Lorentz-Lorentz relationship.

$$\alpha_m = \frac{3R_m}{4\pi N} \quad (11)$$

Where N is Avogadro's number and the molar electronic polarizability is measured in \AA^3 . It is well known that Bi^{3+} cation possesses a very high polarizability (1.508\AA^3), which is due to their large ionic radii and small cation unit field strength. Moreover, Bi^{3+} ions possess a lone pair in the valence shell. This relation presents a general tendency toward an increase in the oxide polarizability with increasing optical basicity.

The refractive index (n) depends on individual ions present in the glass and polarizability of cations. In general, the refractive index of the glass increases for the highly polarizable cations as a result of increase in the non-bridging oxygen to bridging oxygen ratio [13]. It was established by Dimitrov and Saka that there is a general trend by which refractive index increases with the increase in electronic polarizability of oxide ions. Since in the present sample the basicity decreases, which results in the decrease of electronic polarizability, therefore the refractive index also decreases upon the replacement of bismuth with silicate in the glass network. A good agreement is observed between theoretical and observed values as shown in Table.1.

The molar refraction is considered as the sum of the contributions of the cationic refraction and oxygen ionic

refraction. It is related to the structure of the glass which measures the bonding condition in the glass [14]. The molar refraction (R_m) derived by Lorentz and Lorentz relationship.

$$R_m = \left[\frac{n^2 - 1}{n^2 + 2} \right] V_m \quad (12)$$

Where the quantity $(n^2-1)/(n^2+2)$ is called the reflection loss and V_m is the molar volume in cm^3 .

According to Duffy [20] the energy gap is given as

$$E_g = 20 \left(1 - \frac{R_m}{V_m} \right)^2 \quad (13)$$

As the value of n decreases, there is a decrease in the value of R_m/V_m due to which E_g increases. If energy gap is large, λ_p will be small and hence σ decreases.

The optical basicity of an oxidic medium is the average electron donor power of all the oxide atoms comprising the medium. Increasing basicity results in increasing negative charge on the oxygen atom and thus, increasing covalency in the cation-oxygen bonding.

Fig. 1 presents the X-ray diffraction pattern of BBSN-1 with Cu-K α (15.42 nm) line of wavelength at the scanning rate of $2^\circ/\text{min}$ and 2θ was varied from 10° to 80° . The XRD pattern of glass exhibits a broad hump, which shows its amorphous nature.

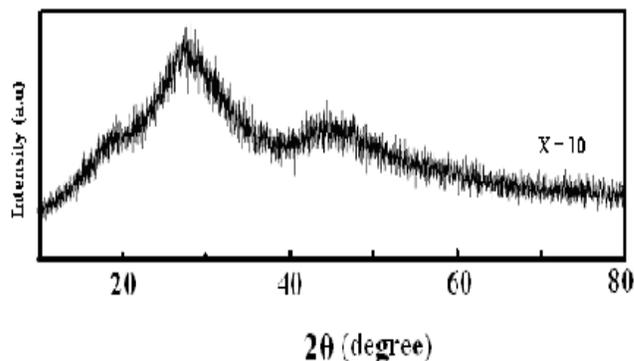


Fig.1. X-ray diffraction pattern of BBSN-1

Table 1: Density (D), molar volume (V_m), refractive index (n), reflection loss (R_L), molar refraction (R_m), electronic polarizability (α_m), energy gap (E_g) and optical basicity (Λ) of $20B_2O_3 \cdot (79.5-x) Bi_2O_3 \cdot xSiO_2 \cdot 0.5 Nd_2O_3$ ($10 \leq x \leq 40$) glasses.

Sample code	x (mol%)	D (g/cm^3)	V_m (cm^3/mol)	n	R_L ($(n^2-1)/n^2+2$)	R_m (cm^3/mol)	α_m (\AA^3)	E_g (eV)	Λ
BBSN1	10	6.92	55.64	1.96	0.486	27.06	10.73	5.28	0.425
BBSN2	20	6.83	49.71	1.87	0.454	22.58	8.95	5.96	0.407
BBSN3	30	6.77	44.00	1.79	0.423	18.59	7.37	6.66	0.387
BBSN4	40	6.64	38.00	1.72	0.395	15.00	5.94	7.32	0.372

Table2:

Oscillator strength for some transitions from the indicated level to the ground level ${}^6\text{H}_{5/2}$, their root mean square (δ_{RMS}) Judd-Ofelt intensity parameters ($\Omega_2, \Omega_4, \Omega_6$) of various Nd^{3+} doped BBSN glasses.

Transitions from ${}^6\text{H}_{5/2} \rightarrow$	$\lambda(\text{nm})$	Oscillator strength (10^{-6})							
		BBSN1		BBSN2		BBSN3		BBSN4	
		f_{expt}	f_{cal}	f_{expt}	f_{cal}	f_{expt}	f_{cal}	f_{expt}	f_{cal}
${}^4\text{G}_{9/2}$	514	1.85	1.96	1.35	1.02	1.22	1.24	1.14	1.12
${}^4\text{G}_{7/2}$	526	3.85	4.32	6.77	5.92	5.35	4.77	5.03	5.23
${}^4\text{G}_{5/2}, {}^2\text{G}_{7/2}$	586	5.53	5.20	4.23	5.10	3.22	3.82	2.35	2.13
${}^4\text{F}_{9/2}$	684	2.99	3.19	1.03	1.07	1.01	0.69	1.06	1.16
${}^4\text{S}_{3/2}, {}^4\text{F}_{7/2}$	750	2.45	2.43	3.53	3.62	3.45	3.51	3.31	3.56
${}^4\text{F}_{5/2}, {}^2\text{H}_{9/2}$	798	2.37	2.23	3.45	3.22	3.16	3.09	3.28	3.34
$\delta_{\text{RMS}} (10^{-6})$		5.07		4.64		3.97		3.69	
$\Omega_2(10^{-20} \text{cm}^2)$		3.52		3.22		2.74		2.40	
$\Omega_4(10^{-20} \text{cm}^2)$		4.19		5.05		5.54		5.59	
$\Omega_6(10^{-20} \text{cm}^2)$		3.86		4.94		5.05		5.07	
Ω_4/Ω_6		1.01		1.02		1.09		1.10	

Table 3:

The peak wavelength (λ_p), radiative transition probability (A_{rad}), branching ratio (β_r), stimulated emission cross-section (σ), total radiative transition probability (A_T) and the radiative life time (τ_r) of Nd^{3+} ion doped BBSN glasses.

Transitions from	BBSN1				BBSN2			BBSN3			BBSN4		
	λ_p (nm)	A_{rad} (s^{-1})	β_r (%)	σ (10^{-20}cm^2)	A_{rad} (s^{-1})	β_r (%)	σ (10^{-20}cm^2)	A_{rad} (s^{-1})	β_r (%)	σ (10^{-20}cm^2)	A_{rad} (s^{-1})	β_r (%)	σ (10^{-20}cm^2)
${}^4\text{F}_{3/2}$	1075	1130	0.536	2.20	1023	0.545	2.03	946	0.543	1.88	789	0.570	1.69
${}^4\text{I}_{11/2}$	905	980	0.464	1.67	854	0.454	1.44	746	0.428	1.07	595	0.430	0.91
$A_T(\text{s}^{-1})$		2110			1877			1742			1384		
$\tau_r(\mu\text{s})$		0.474			0.532			0.574			0.722		

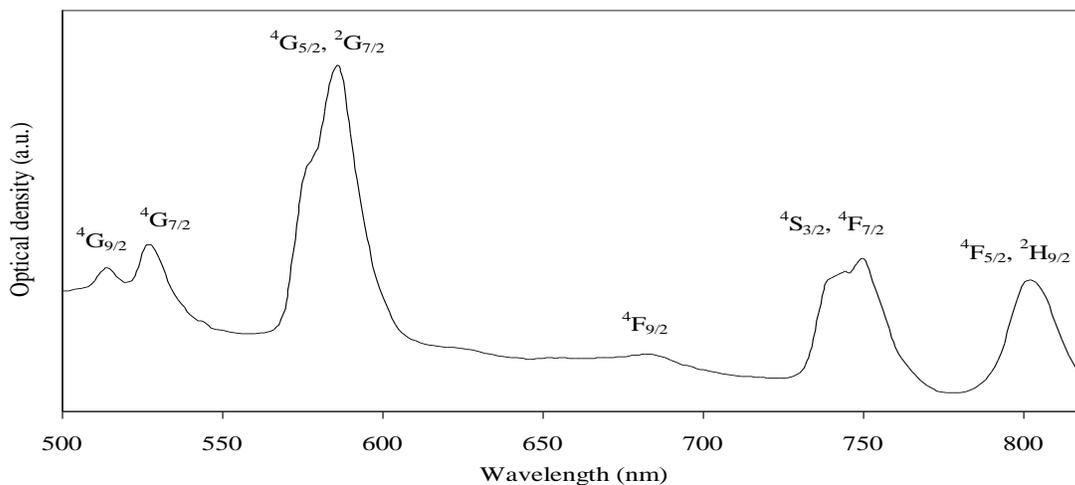


Fig. 2. Absorption spectrum of $20\text{B}_2\text{O}_3, 69.5\text{Bi}_2\text{O}_3, 10\text{SiO}_2, 0.5 \text{Nd}_2\text{O}_3$. (BBSN-1) glass sample.

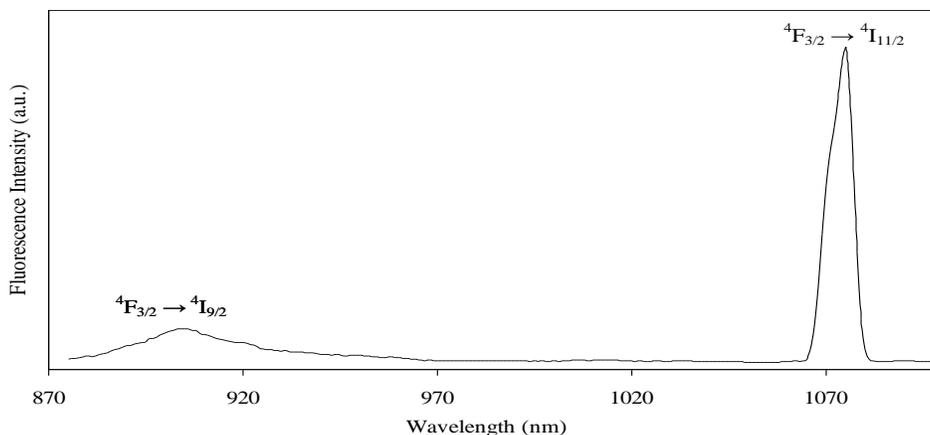


Fig. 3. Emission spectra of 20B₂O₃ 69.5Bi₂O₃ 10SiO₂ 0.5 Nd₂O₃. (BBSN-1) glass sample.

4.2 Absorption spectra analysis

The optical absorption spectrum of BBSN-1 glass sample in the wavelength range 500-850 nm is shown in Fig.2. The absorption spectra for all other samples are qualitatively similar. From Fig. 2, the absorption bands at 514, 528, 586, 684, 750, 802 nm are observed, which predict the Nd³⁺ ion transitions from the ground state ⁴I_{9/2} to various excited state within the 4f shell and are assigned as ⁴I_{9/2} → ⁴G_{9/2}, ⁴G_{7/2}, ⁴G_{5/2+2}G_{7/2}, ⁴F_{9/2}, ⁴S_{3/2}, ⁴F_{5/2+2}H_{9/2}, respectively. It is worth noting that the transition ⁴I_{9/2} → ⁴G_{5/2+2}G_{7/2} is the most intense than other transitions.

The values of oscillator strength obtained from experimental spectra and those calculated using the Judd-Ofelt parameters are given in Table 2 along with the respective root mean square (δ_{RMS}) deviations. From Table 2 it is observed that with the addition of SiO₂ in the host glasses the values of f_{expt}, f_{cal} and δ_{RMS} decreases. According to the Judd-Ofelt theory [11-12], the effect of the crystal field can be described by three intensity parameters Ω_λ (λ= 2, 4, 6). The intensity parameters are obtained from the experimental oscillator strengths and the calculated double reduced matrix elements by using least square analysis method [15].

The matrix elements used in the fitting procedure were those given by Carnall et al. [16]. The intensity parameters determined for Nd³⁺ ions doped glasses are listed in Table 3. The intensity parameters are found to be in the order Ω₄ > Ω₆ > Ω₂ for all the glass samples. These intensity parameters reflect the local structure and bonding in the vicinity of rare earth ion to some extent. The parameter Ω₂ is associated with asymmetry of the ligand field near the rare earth ion so its value is expected to be higher when the symmetry of the ligand field at the rare earth site is lower. The decrease of Ω₂ with decrease in Bi₂O₃ content for Nd³⁺ ions indicate that the rare earth ion in lead bismuth silicate glasses might be surrounded by bismuthate groups that affect the symmetry of ligand field at the rare earth ion site.

The intensity parameter Ω₆ is related to covalency of RE-O bond and decreases with increasing covalency

where as covalency of the RE-O bond is assumed to be related with the local basicity around the rare earth sites [17]. The optical basicity of Bi₂O₃ is higher than that of SiO₂, Therefore with decrease in Bi₂O₃ content in the host glass, covalency of Nd-O bond increases which results in the decrease in Ω₆ values. As a general conclusion, Ω₂ and Ω₆ parameters decrease with the increase in the degree of covalency of the rare earth oxygen bonds.

The intensity parameter Ω₄ is related to the rigidity of the host glass [18]. In the present glass samples the rigidity of the host glass decreases with the decrease in Bi₂O₃ content (Table3), similar results were reported earlier in Er³⁺ ion doped zinc bismuth borate glass [19].

4.3 Fluorescence spectra analysis

Besides the main emission, near 1060 nm (⁴F_{3/2} → ⁴I_{11/2}) another transition at 905 nm (⁴F_{3/2} → ⁴I_{9/2}) is observed at λ_{ext} = 800 nm as shown in Fig.3. The ratio Ω₄ / Ω₆ which is usually considered as the spectroscopic quality factor (SQF) can be used to describe the strength of the emission transitions. The emission cross-section is expected to increase with increasing refractive index of a glass host, because the emission cross-section that is due to the electric dipole transitions of rare earth ions increases as the refractive index of the glass host [σ_e ~ (n²+2) / n] increases.

For laser applications, the value of emission cross-section is an important parameter and signifies the rate of energy extraction from the optical material. A large stimulated emission cross-section is of benefit for a low threshold and a high gain in laser operation [20]. The effective bandwidth of the emission spectra is mainly caused by splitting of the levels of transitions and the site to site variation of the ligand field around Nd³⁺ ions in the glass i.e., inhomogeneous broadening. The slight decrease in effective bandwidth implies that the asymmetry of the ligand field becomes weaker. This result is again in accordance with decrease in the value of Ω₂ with Bi₂O₃ content in the host glass. From Table 3, it is observed that the stimulated emission cross-section (σ_e) and fluorescence intensity decrease with the decrease in Bi₂O₃ content in the host glass. These variations are in agreement with previous investigations. The high values of stimulated emission cross-section suggest that the prepared

glasses are potential candidates for the lasing host materials and for low threshold, high gain applications.

V. CONCLUSIONS

The refractive index of the glass reduces from 1.96 to 1.72 with the decrease in $\text{Bi}_2\text{O}_3 / \text{SiO}_2$ ratio. The spectroscopic properties of Nd^{3+} ions doped glasses have been analyzed on the basis of Judd-Ofelt theory. The intensity parameter Ω_2 is most sensitive to the local environment of Nd^{3+} ions, larger values ($\Omega_2 > 2$) suggest a less Centro symmetric coordination environment. The branching ratio for the ${}^4\text{F}_{3/2} \rightarrow {}^4\text{I}_{11/2}$ transition in the present glass is 57% and the predicted spontaneous radiative transition rate is as high as 2110 s^{-1} for BBSN1 sample. The stimulated emission cross-section for the ${}^4\text{F}_{3/2} \rightarrow {}^4\text{I}_{11/2}$ transition decreases from 2.20×10^{-20} to $1.69 \times 10^{-20} \text{ cm}^2$, and it is comparable to the some of frequently used laser glasses. The results of these investigations indicate that neodymium doped bismuth boro-silicate glasses may be suitable host for the lasing materials. The IR absorbance analysis of these glasses confirmed the decrease in both the NBOs and the molar volume, which in turn explain the obtained decrease in the refractive indices, the polarizability, and the optical basicity of the studied glasses. Also, the observed decrease in the optical band gap is related to the weaker bond strength of the Bi-O compared to the Si-O.

REFERENCES

- [1] A.A. Kamniskii Laser crystals. second ed. Springer, Berlin (1990), pp. 252-253
- [2] E. Pacoraro, J.A. Sampaio, L.A.O. Nunes, S Gama and M.L. Baesso. Spectroscopic properties of Water free Nd_2O_3 Doped Low Silica Aluminosilicate Glass. Non-Cryst. Solids 77(2000)[3]B. Karthikeyan, R. Philip, S. Mohan. Optical and Non-linear Optical Properties of Nd^{3+} doped Heavy Metal Borate Glasses. Optics Commun 246 (2005) 153.
- [4] Boehm L., Reisfeld R., Sepctor N., Optical transitions of Sm^{3+} in oxide glasses, Journal of Solid State Chemistry, 28 (1979) 75.
- [5] Stone B.T., Bray K.L., Fluorescence properties of Er^{3+} -doped sol-gel glasses, Journal of Non-Crystalline Solids, 197 (1996) 136.
- [6] Reisfeld R., Jorgensen C.K., Lasers and Excited States of Rare Earths, Springer, Berlin, (1977).
- [7] Ratnakaram Y.C., Thirpathi N.D., Chakaradhar R.P.S., Spectral studies of Sm^{3+} and Dy^{3+} doped lithium cesium mixed alkali borate glasses, Journal of Non-Crystalline Solids, 352 (2006) 3914.
- [8] Jiangting Sun, Jihua Zhang, Baojiu Chen et al. Preparation and optical properties of Er^{3+} doped gadolinium borosilicate glasses. Journal of Rare Earths 23(2) (2005) 157.
- [9] I. Operea, H. Hesse, K. Betzler. Optical properties of bismuth borate glasses. Optical Materials 26(3) (2004) 235.
- [10] P. Becker. Thermal and Optical Properties of Glasses of the System $\text{Bi}_2\text{O}_3\text{-B}_2\text{O}_3$ Cryst. Res. Technol. 1 (2003) 74
- [11] B.R. Judd. Optical Absorption Intensities of Rare-Earth Ions. Phys. Rev. 127 (1962) 750.
- [12] J.S. Ofelt. Intensities of Crystal Spectra of Rare-Earth Ions. Chem Phys, 37 (1962) 511.
- [13] X. Zou, T. Izumitani Spectroscopic properties and mechanisms of excited state absorption and energy transfer upconversion for Er^{3+} doped glasses. J. Non-Cryst. Solids 162 (1) (1993) 68.
- [14] S. Tanabe, T Ohayagi and N Soga, et al. Compositional dependence of Judd-Ofelt parameters of Er^{3+} ions in alkali metal borate glass. Phys. Rev. B46 (1992) 3305.
- [15] W.F. Krupke, Induced-emission cross-section in neodymium laser glasses. IEEE J. Quantum electron QE-10 (1974) 450-457
- [16] W.T. Carnall, P.R. Fdcids, K. Rajnax. Electronic energy levels in the trivalent lanthanide aquo ions. The journals of Chemical Physics 49(10) (1968) 4424.
- [17] A. Agarwal, I. Pal, S. Sanghi, M.P. Aggarwal Judd-Ofelt Parameter and Radiative Properties of Sm^{3+} ions Doped Zinc Bismuth Borate Glasses”, Opt. Mater, 32 (2009) 399.
- [18] I Pal, A. Agarwal, S. Sanghi, M.P. Aggarwal. Structural, Absorption and Flurescence Spectral Analysis of Pr^{3+} ions Doped Zinc Bismuth Borate Glasses. Alloys Compds, 509 (2011) 7625.
- [19] S. Sanghi, I. Pal, A. Agarwal, M.P. Aggarwal. Effect of Bi_2O_3 on Spectroscopic and Structural Properties of Er^{3+} Doped Cadmium Bismuth Borate Glasses, Spectro. Chema Acta 83 (2011) 94.
- [20] M.J. Weber. Probabilities for radiative and nonradiative decay of Er^{3+} in LaF_3 . Physical Review 157(2) (1967) 157.

A Comparative Study of Combination of Different Bit Positions In Image Steganography

Ramanpreet Kaur¹, Baljit Singh², Ishpreet Singh³

¹(Research scholar of Department of CSE and IT, BBSBEC Fatehgarh Sahib, Punjab, India)

^{2,3}(Associate Prof of .Department of CSE and IT, BBSBEC Fatehgarh Sahib, Punjab, India)

ABSTRACT: Steganography hides the existence of the data inside any cover file. There are different file formats used in steganography like text, image, audio and video. Out of these file formats image steganography is followed in this paper. One of the major objective of hiding data using image steganography is to hide the data in an image, so that the changes in the intensity of the colors of image must not be visible to, human eye. The focus of this paper is on spatial domain technique i.e. LSB technique of image steganography. Method used in the paper hides the data in combination of LSBs instead of hiding the data only in least significant one bit. Combination of bits used are LSB (1, 2) bits and (2, 3) bits. Results are compared qualitatively and quantitatively using parameters PSNR, MSE, BER, Entropy, Standard deviation.

Keywords: BER, Entropy, Image steganography, LSB technique, MSE, PSNR, Std. deviation.

I. INTRODUCTION

Due to the increase in the use of internet it becomes important to secure the sensitive data and information on internet. Therefore, to secure the sensitive information on the internet various techniques like cryptography, steganography are used to hide the data in digital media. Cryptography only keeps the contents of the message secret but sometimes it is necessary to keep the existence of the message secret. So, a technique which keeps the existence of a message secret is known as steganography [1].

This paper explores the steganography technique for security purpose. Paper presents the LSB technique of image steganography in which data is hidden in the combination of LSBs of image pixels.

II. STEGANOGRAPHY

1. Steganography

Steganography is a method of secret communication that inserts the message inside any cover in the way that third party should not be able to find it out that a message is hidden in the cover. The cover can be an image, text, audio or video. The word steganography is derived from the Greek word 'Steganos', which means covered or secret and 'graphy' means writing or drawing, that combined means, "Covered Writing" or "Secret Writing".

The goal of steganography is to hide the sensitive information in a cover so that nobody can guess the existence of it. Information hiding technique becomes important in a number of application areas [2, 3]. The growing need of communications through the internet demands the confidentiality and integrity of data to protect against unauthorized access. Therefore, for the security of

data from unauthorized users there is a need of hiding the data. There are various techniques of steganography of which image steganography technique is most useful.

2. Image Steganography

An image is an array of M*N matrix. Each and every pixel has a numerical value which represents the color and light intensity of the pixel [4]. Images are more popular cover files for transmission over the internet due to harmlessness and attraction. In image steganography, data is to be inserted into the cover image that gives the resultant stego-image [5, 6].

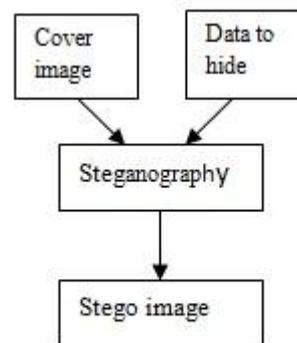


Fig.1. Process of hiding data in image.

There are various steganographic techniques for image file formats are classified as [2]:

1. Spatial domain technique.
2. Masking and filtering.
3. Transform techniques.

Spatial domain technique embeds the message directly in the intensity of the pixels. Spatial domain technique includes least significant bit (LSB).

3. LSB Technique

This is one of the simplest and easiest methods of hiding data in images. In this technique, the data in binary form is to be hidden into the LSBs of the carrier bytes or in pixels of image. The overall change to the image is so small that human eye would not be able to discover. In 24-bit images each 8-bit value refers to the red, green and blue color. But in 8-bit images each pixel is of 8-bits, so each pixel stores maximum 256 colors [7, 8].

For example, in our method to hide a letter A whose binary value is 10000001, in the combination of LSBs of an 8-bit image then we need four pixels.

Suppose the original four pixels are:
00100101
11100101

11001011
 01000111
 Secret data i.e. A (10000001)
 The resulting four pixels are as follows:
 00100101
 11100100
 11001000
 01000110

Only 4 bits needed to be changed to embed the secret data into the first 4 pixels of the image. On average, LSB requires that only half the bits in an image be changed to hide a secret message [9, 10, 11].

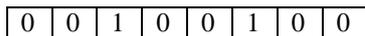
III. Proposed Work

The proposed work includes the hiding of data in combination of two LSBs. The selected pair of bits (1, 2) or (2, 3) of pixels in image has been replaced with the data bits. The cover image selected is of grayscale.

Procedure:

1. Extract all the characters of secret data and store them in character array.
2. Replace the pair of bits (1, 2) or (2, 3) of selected pixel with length of data and bits of characters of Character-Array.

Suppose to hide bits of character 8 bits of the first pixel are:



To hide character A whose binary value is 10000001 and 2-bits of A are:



3. Repeat steps till all the characters have been embedded.
4. Obtained image will hide all the characters that we input.
5. Compare pair bits (1, 2) and (2, 3).

We have used performance parameters like PSNR, MSE, BER, entropy, std. deviation to check the impact on image, of replacement of different pair bits.

PSNR stands for peak signal to noise ratio. It used to measure the quality of images. Higher the PSNR value, better the quality of image. It is estimated in decibels (db). It is defined as:

$$PSNR = 10 \log_{10} \frac{R^2}{MSE} \quad (1)$$

In grayscale images R is 255.

MSE stands for the mean square error. Lower the value of MSE, better the quality of image. It is defined as:

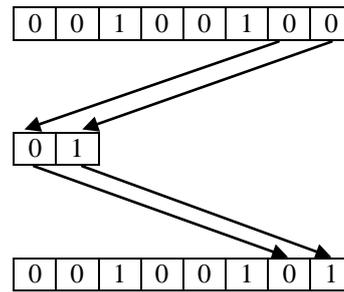
$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - y_{ij})^2 \quad (2)$$

Here M and N represents the no. of rows and columns of the image respectively. x_{ij} is the original image and y_{ij} is the stego image.

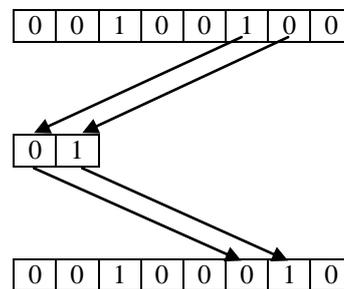
BER is bit error rate. It is explained as:

$$BER = 1/PSNR \quad (3)$$

To replace the pair of bits (1, 2) of first pixel with the data bits:



Similarly, to replace the pair of bits (2, 3) of first pixel with the 2Bits of A are:



Repeat the same process till all the bits of A has been embedded.

Entropy is used to measure the randomness of image. It is defined by the following equation:

$$Entropy = - \sum_{i=1}^l P_i \log_2(P_i) \quad (4)$$

Where P_i is the probability of getting a particular intensity and l is the total intensity values.

Standard deviation is defined by the equation:
 Standard deviation

$$\sigma_j = \left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (5)$$

Where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

IV. EXPERIMENTAL RESULTS

Experimental results evaluate the performance by hiding the data in LSB bit pairs of pixels in gray scale images using MATLAB. Various performance parameters like PSNR, MSE, BER, entropy, standard deviation have been used to evaluate the performance. The experiment has been taken out on various images which includes both standard and natural images by embedding ASCII characters. Figure 2-10 shows original images, stego-images with data hiding on 1, 2 bit pair and stego-images with data hiding on 2, 3 bit pair. Figures from 11 to 28 shows PSNR, BER, MSE graphs of different images. Table 1 and Table 2 shows the results of three images

Lena, Autumn and Neptune where quality factor is 255. Table 1 shows performance of images on bit pair (1, 2) and Table 2 shows the performance of images on bit pair (2, 3).



Fig.2. Original Lena



Fig.3. Original Autumn



Fig.4. Original Neptune



Fig.5. Stego-image (1, 2)



Fig.6. Stego-image (1, 2)



Fig.7. Stego-image (1, 2)

Fig.8. Stego-image (2, 3)

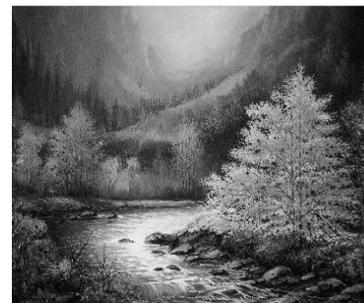


Fig.9. Stego-image (2, 3)



Fig.10. Stego-image (2, 3)

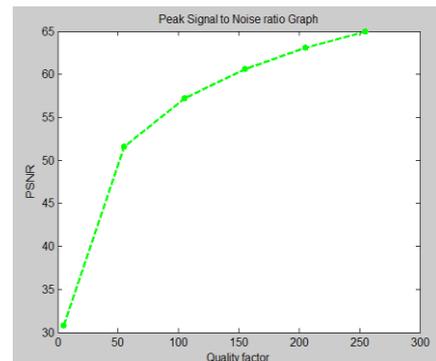


Fig.11.PSNR of Lena (1, 2)

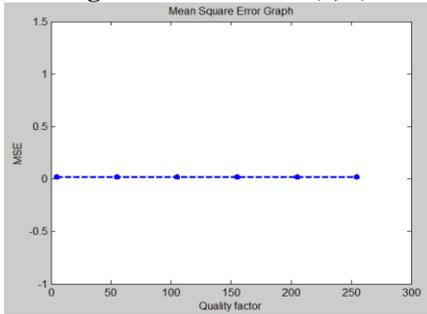


Fig.12.MSE of Lena (1, 2)

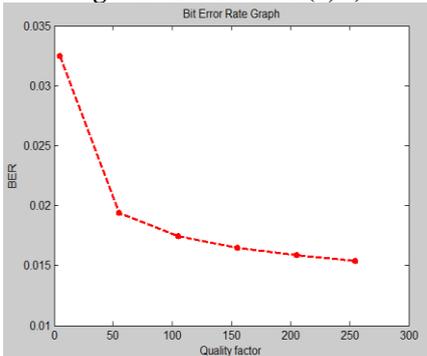


Fig.13.BER of Lena (1, 2)

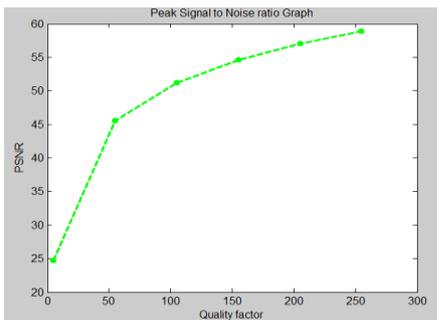


Fig.14.PSNR of Lena (2, 3)

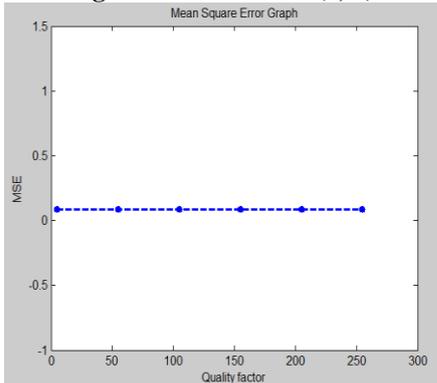


Fig.15.MSE of Lena (2, 3)

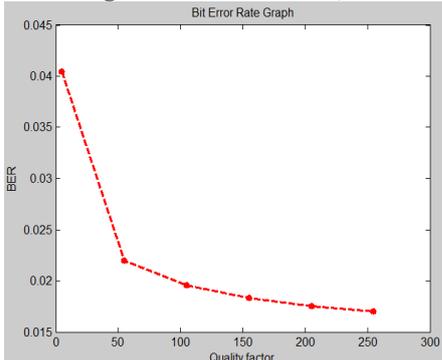


Fig.16.BER of Lena (2, 3)

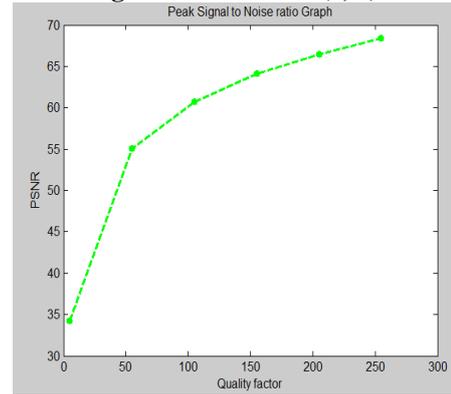


Fig.17.PSNR of Autumn (1, 2)

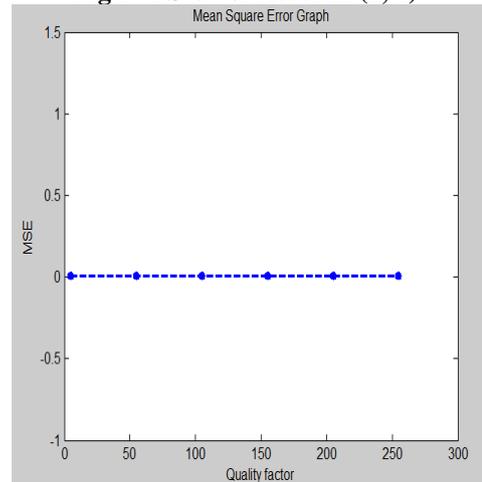


Fig.18.MSE of Autumn (1, 2)

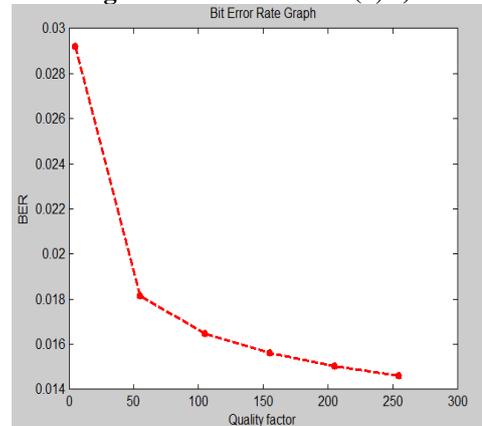


Fig.19.BER of Autumn (1, 2)

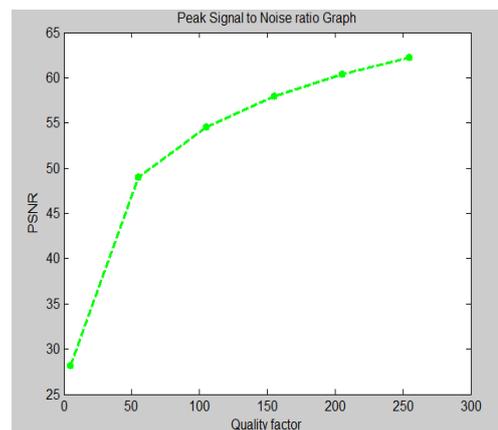


Fig.20.PSNR of Autumn (2, 3)

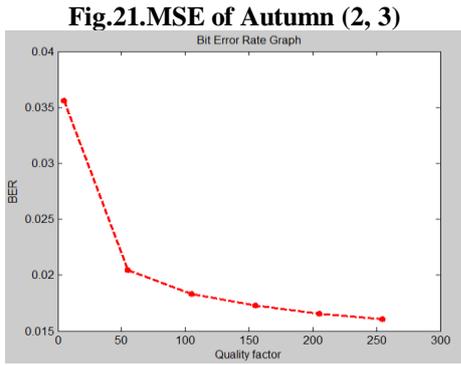


Fig.21.MSE of Autumn (2, 3)

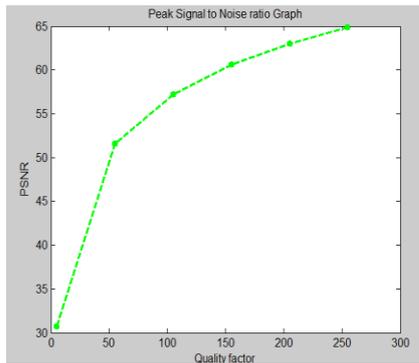
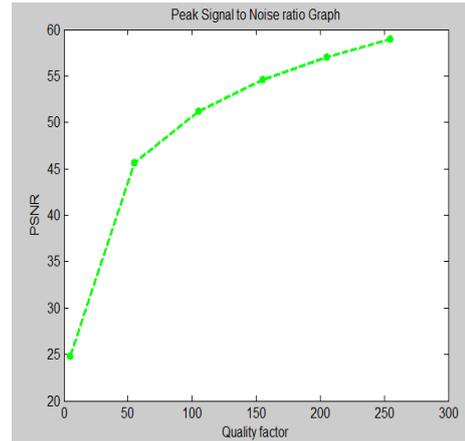


Fig.23.PSNR of Neptune (1, 2)

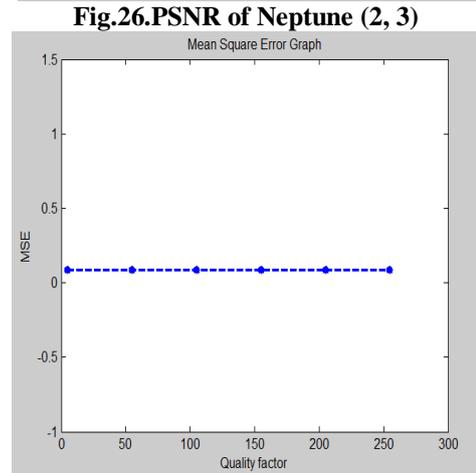


Fig.24.MSE of Neptune (1, 2)

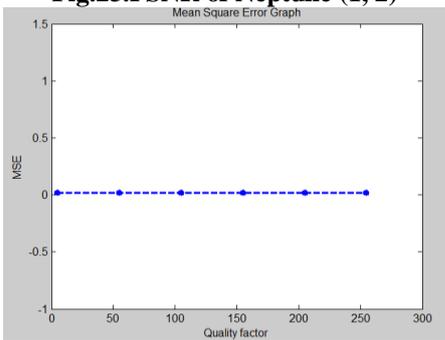


Fig.25.BER of Neptune (1, 2)

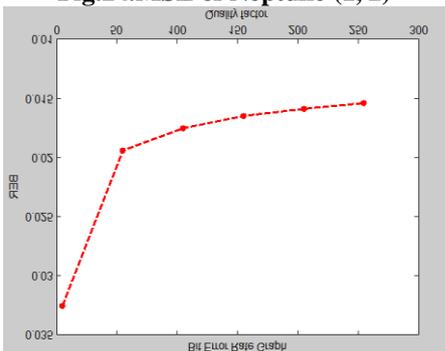


Fig.26.PSNR of Neptune (2, 3)

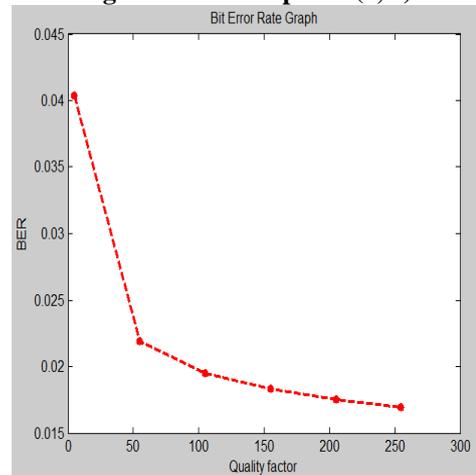


Fig.27.MSE of Neptune (2, 3)

Table.1. Results of (1, 2) bit

Parameters	Images		
	Lena	Autumn	Neptune
PSNR	58.8896	62.2717	58.9339
BER	0.0169809	0.0160587	0.0169682
MSE	0.0839691	0.0385403	0.0831168
Entropy	7.44642	7.65937	7.52104
Std. Dev.	12.5371	9.97252	13.8436

Table.2. Results of (2, 3) bit

Parameters	Images		
	Lena	Autumn	Neptune
PSNR	64.9506	68.3888	64.9057
BER	0.0153963	0.0146223	0.015407
MSE	0.0207977	0.00942315	0.0210141
Entropy	7.44612	7.65929	7.52059
Std. Dev.	12.5384	9.97278	13.842

V. Conclusion

This paper shows the results of comparison of hiding the data in LSB (1, 2) bit pair and in (2, 3) bit pair. It is concluded that hiding the data in the (1, 2) bit pair shows better results than hiding the data in (2, 3) bit pair. Test is made on various images which includes both standard and natural images of which three are shown, using more than one parameter. The parameter implies that hiding the secret message in (1, 2) bit pair is less visible to human eye i.e. quality of image is better as compare to the second one. In future, we can also extend this method for more combination of bit positions and this can also be performed on the color images.

References

[1]. T. Morkel, J.H.P. Eloff, M.S. Olivier, 2005. "An Overview of Image Steganography", in Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa.

[2]. Muhalim Mohamed Amin, Subariah Ibrahim, Mazleena Salleh and Mohd Rozi Katmin, 2003. "Information Hiding using Stegnography", Department of Computer System & Communication Faculty of Computer Science and Information system, universiti teknologi malaysia.

[3]. Arvind Kumar and Km. Pooja, 2010. "Steganography- A Data Hiding Technique", International Journal of Computer Applications, Vol. 9, No.7.

[4]. Neil F. Johnson, Sushil Jajodia 1998. "Exploring Steganography: Seeing the Unseen", Dept. of Information and Software Systems Engineering, George Mason University, Fairfax.

[5]. Jagvinder Kaur and Sanjeev Kumar, 2011 "Study and Analysis of Various Image Steganography Techniques", International Journal of Computer Science and Technology (IJCSST), ISSN: 2229-4333(Print)|ISSN:0976-491(Online), Vol.2, Issue3.

[6]. Kaveh Ahmadi, 2011. "A New Method for Image Security and Data Hiding in Image", American Journal of Scientific Research ISSN 1450-223X Issue 38(2011), pp. 41-49.

[7]. Beenish Mehboob and Rashid Aziz Faruqi, 2008. "A Stegnography Implementation", Department of Computer Science and Engineering, Bahria University, Karachi, Pakistan, 1-4244-2427-6/08/\$20.00 ©2008 IEEE.

[8]. Jassim Mohmmmed Ahmed and Zulkarnain Md Ali, 2011. "Information Hiding using LSB technique", IJCSNS International 18 Journal of Computer Science and Network Security, Vol.11, No.4.

[9]. A.E.Mustafa, A.M.F.ElGamal, M.E.ElAlmi and Ahmed.BD, 2011. "A Proposed Algorithm For Steganography In Digital Image Based on Least Significant Bit", Research Journal Specific Education Faculty of Specific Education, Mansoura University, Issue No. 21.

[10]. Vijay Kumar Sharma and Vishal Shrivastava, 2012. "A Steganography Algorithm for hiding image in image by improved LSB substitution by minimize detection", Journal of Theoretical and Applied Information Technology, ISSN: 1992-8645, Vol. 36, No.1.

[11]. Amanpreet Kaur, Renu Dhir, and Geeta Sikka, 2009. "A New Image Steganography Based On First Component Alteration Technique", International Journal of Computer Science and Information Security (IJCSIS), Vol.6, No.3.

Data Caching Between Mobile Nodes in Wireless Adhoc Networks

Y.Tulasi Rami Reddy 1, K.Gopinath 2

¹ M.Tech Student In CSE Department, Kottam College Of Engineering, Kurnool (India) 518218,

² Associate Professor in IT Department, KSRM College Engineering, Kadapa (India) 516003

Abstract: we are introduced the cooperative caching in wireless networks ,where the nodes may be mobile and exchange information in apeer-to-peer fashion. We consider both cases of nodes with large-and small-sized caches. For large-sized caches, we devise a strategy where nodes, independent of each other, decide whether to cache some content and for how long. In the case of small-sized caches, we aim to design a content replacement strategy that allows nodes to successfully store newly received information while maintaining the good performance of the content distribution system. Under both conditions, each node takes decisions according to its per-ception of what nearby users may store in their caches and with the aim of differentiating its own cache content from the other nodes'. The result is the creation of content diversity within the nodes neighborhood so that a requesting user likely finds the de-sired information nearby. We simulate our caching algorithms indifferent ad hoc network scenarios and compare them with other caching schemes, showing that our solution succeeds in creating the desired content diversity, thus leading to a resource-efficient Information access.

Index Terms – Data cashing, mobile Nodes, mobile ad hoc networks.

I. Introduction

The necessary information to users on the move is one of the most promising directions of the infotainment business, which rapidly becomes a market reality, because infotainment modules are deployed on cars and handheld devices. The ubiquity and ease of access of third- and fourth-generation (3G or 4G) networks will encourage users to constantly look for content that matches their interests. However, by exclusively relying on downloading from the infrastructure, novel applications such as mobile multimedia are likely to overload the wireless network (as recently happened to AT&T following the introduction of the iPhone [1]). It is thus conceivable that a peer-to-peer system could come in handy, if used in conjunction with cellular networks, to promote content sharing using ad hoc networking among mobile users [2]. For

highly popular content, peer-to-peer distribution can, indeed, remove bottlenecks by pushing the distribution from the core to the edge of the network. In such an environment, however, a cache-all-you-see approach is unfeasible, because it would swamp node storage capacity with needless data that were picked up on the go. Thus, several techniques of efficiently caching information in wireless ad hoc networks have been investigated in the literature; for example, see the surveys in [3] and [4] and the related work discussed in Section II. The solution that we propose, called Hamlet, aims at creating content diversity within the node neighborhood so that users likely find a copy of the different information items nearby (regardless of the content popularity level) and avoid flooding the network with query messages. Although a similar concept has been put forward in [5]–[8], the novelty in our proposal resides in the probabilistic estimate, run by each node, of the information presence (i.e., of the cached content) in the node

proximity. The estimate is performed in a cross-layer fashion by overhearing content query and information reply messages due to the broadcast nature of the wireless channel. By leveraging such a local estimate, nodes autonomously decide what information to keep and for how long, resulting in a distributed scheme that does not require additional control messages. The Hamlet approach applies to the following cases.

1.1 Large-sized caches. In this case, nodes can potentially store a large portion (i.e., up to 50%) of the available information items. Reduced memory usage is a desirable (if not required) condition, because the same memory may be shared by different services and applications that run at nodes.

1.2 Small-sized caches. In this case, nodes have a dedicated but limited amount of memory where to store a small percentage (i.e., up to 10%) of the data that they retrieve. The caching decision translates into a cache replacement strategy that selects the information items to be dropped among the information items just received and

the information items that already fill up the dedicated memory.

We evaluate the performance of Hamlet in different mobile network scenarios, where nodes communicate through ad hoc connectivity. The results show that our solution ensures a high query resolution ratio while maintaining the traffic load very low, even for scarcely popular content, and consistently along different network connectivity and mobility scenarios.

II. RELATED WORK

Several papers have addressed content caching and content replacement in wireless networks. In the following sections, we review the works that are most related to this paper, highlighting the differences with respect to the Hamlet framework that we propose.

2.1. Cooperative Caching

In [9], distributed caching strategies for ad hoc networks are presented according to which nodes may cache highly popular content that passes by or record the data path and use it to redirect future requests. Among the schemes presented in [9], the approach called HybridCache best matches the operation it as a benchmark for Hamlet in our comparative evaluation. In [10], a cooperative caching technique is presented and shown to provide better performance than HybridCache. However, the solution that was proposed is based on the formation of an over-layer network composed of “mediator” nodes, and it is only fitted to static connected networks with stable links among nodes. These assumptions, along with the significant communication overhead needed to elect “mediator” nodes, make this scheme unsuitable for the mobile environments that we address. The work in [11] proposes a complete framework for information retrieval and caching in mobile ad hoc networks, and it is built on an underlying routing protocol and requires the manual setting of a network wide “cooperation zone” parameter. Note that assuming the presence of a routing protocol can prevent the adoption of the scheme in [11] in highly mobile networks, where maintaining network connectivity is either impossible or more communication expensive than the querying/caching process. Furthermore, the need of a manual calibration of the “cooperation zone” makes the scheme hard to configure, because different environments are considered. Conversely, Hamlet is self contained and is designed to self adapt to network environments with different mobility and connectivity features.

2.2. Content Diversity

Similar to Hamlet, in [6], mobile nodes cache data items other than their neighbors to improve data accessibility. In particular, the solution in [6] aims at caching copies of the same content farther than a given number of hops. Such a scheme, however, requires the maintenance of a consistent state among nodes and is unsuitable for mobile network topologies. The concept of caching different content within a neighborhood is also exploited in [7], where nodes with similar interests and mobility patterns are grouped together to improve the cache hit rate, and in [8], where neighboring mobile nodes implement a cooperative cache replacement strategy. In both works, the caching management is based on instantaneous feedback from the neighboring nodes, which requires additional messages. The estimation of the content presence that we propose, instead, avoids such communication overhead.

2.3. Caching With Limited Storage Capability

In the presence of small-sized caches, a cache replacement technique needs to be implemented. Aside from the scheme in [8], centralized and distributed solutions to the cache placement problem, which aim at minimizing data access costs when network nodes have limited storage capacity, are presented in [14]. Although centralized solutions are not feasible in ad hoc environments, the distributed scheme in [14] makes use of cache tables, which, in mobile networks, need to be maintained similar to routing tables. A content replacement scheme that addresses storage limitations is also proposed in [16]. It employs a variant of the last recently used (LRU) technique, which favors the storage of the most popular items instead of the uniform content distribution targeted by Hamlet. In addition, it exploits the cached item IDs provided by the middleware to decide on whether to reply to passing-by queries at the network layer, as well as link-layer traffic monitoring to trigger prefetching and caching. In [17], the popularity of content is taken into account, along with its update rate, so that items that are more frequently updated are more likely to be discarded. Similarly, in [18], cache replacement is driven by several factors, including access probability, update frequency, and retrieval delay.

2.4. Data Replication

Although addressing a different problem, some approaches to data replication are relevant to the data caching solution that we propose. One technique of eliminating information replicas among neighboring nodes

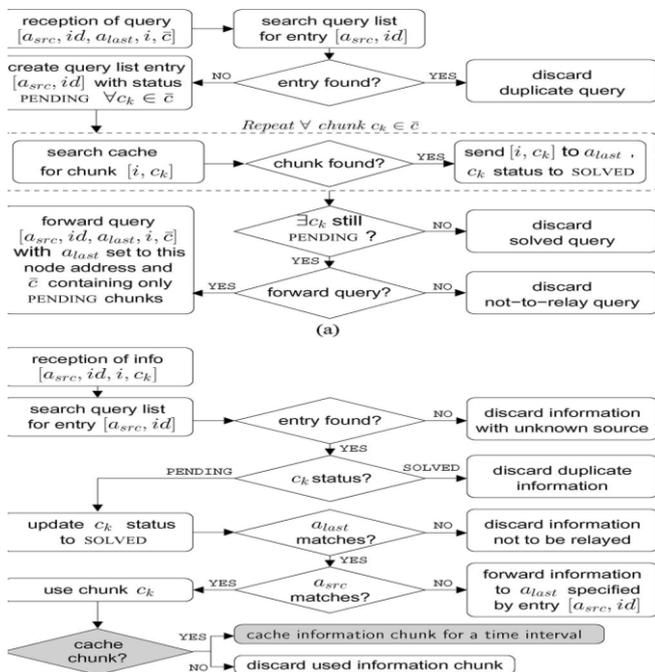
is introduced in [21], which, unlike Hamlet, requires knowledge of the information access frequency and periodic transmission of control messages to coordinate the nodes' caching decisions. In [5], the authors propose a replication scheme that aims at having every node close to a copy of the information and analyze its convergence time.

III. System Outline And Assumptions

Hamlet is a fully distributed caching strategy for wireless ad hoc networks whose nodes exchange information items in a peer-to-peer fashion. In particular, we address a mobile ad hoc network whose nodes may be resource-constrained devices, pedestrian users, or vehicles on city roads. Each node runs an application to request and, possibly, cache desired information items. Nodes in the network retrieve information items from other users that temporarily cache (part of) the requested items or from one or more gateway nodes, which can store content or quickly fetch it from the Internet. We assume a content distribution system where the following assumptions hold: 1) A number I of information items is available to the users, with each item divided into a number C of chunks; 2) user nodes can overhear queries for content and relative responses within their radio proximity by exploiting the broadcast nature of the wireless medium; and 3) user nodes can estimate their distance in hops from the query source and the responding node due to a hop-count field in the messages.

1

Fig: 1. Flow Charts of the processing of a) query and b) information messages at user nodes



Although Hamlet can work with any system that satisfies the aforementioned three generic assumptions, for concreteness, we detail the features of the specific content retrieval system that we will consider in the remainder of this paper. The reference system that we assume allows user applications to request an information item i ($1 \leq i \leq I$) that is not in their cache. Upon a request generation, the node broadcasts a query message for the C chunks of the information item. Queries for still missing chunks are periodically issued until either the information item is fully retrieved or a timeout expires. If a node receives a fresh query that contains a request for information i 's chunks and it caches a copy of one or more of the requested chunks, it sends them back to the requesting node through information messages. If the node does not cache (all of) the requested chunks, it can rebroadcast a query for the missing chunks, thus acting as a forwarder. Once created, an information message is sent back to the query source. To avoid a proliferation of information copies along the path, the only node that is entitled to cache a new copy of the information is the node that issued the query. Information messages are transmitted back to the source of the request in a unicast fashion, along the same path from which the request came.

A node that receives the requested information has the option to cache the received content and thus become a provider for that content to the other nodes.

IV. Simulation Scenarios And Metrics

We tested the performance of Hamlet through ns2 simulations under the following three different wireless scenarios:

1) a network of vehicles that travel in a city section (referred to as City); 2) a network of portable devices carried by customers who walk in a mall (Mall); and 3) a network of densely and randomly deployed nodes with memory limitations (memory constrained nodes). The three scenarios are characterized by different levels of node mobility and network connectivity. In the City scenario, as depicted in Fig. 4, vehicle movement is modeled by the intelligent driver model with intersection management (IDM-IM), which takes into account car-to-car interactions and stop signs or traffic lights [27]. We simulated a rather sparse traffic, with an average vehicle density of 15 veh/km over a neighborhood of 6.25 km². The mobility model settings, forcing vehicles to stop and queue at intersections, led to an average vehicle speed of about 7 m/s (i.e., 25 km/h). We set the radio range to 100 m in the vehicular scenario, and by analyzing the network topology during the simulations, we observed an average link duration of 24.7 s

and a mean of 45 disconnected node clusters concurrently present over the road topology. The City scenario is thus characterized by scattered connectivity and high node mobility. The Mall scenario is represented in Fig. 4 as a large L-shaped open space of 400 m of length on the long side, where pedestrian users can freely walk. In this scenario, we record an average of 128 users who walk at an average speed of 0.5 m/s according to the random-direction mobility model with reflections [28]. The node radio range is set to 25 m, leading to an average link duration equal to 43 s, with a mean of ten disconnected clusters of users present at the same time in the network. The connectivity level in the Mall is thus significantly higher than in the City, whereas node mobility is much lower. The memory-constrained scenario is similar to the scenario employed for the performance evaluation of the cache. The information-sharing application lies on top of a User Datagram Protocol (UDP)-like transport protocol, whereas,

at the media access control (MAC) layer, the IEEE 802.11 standard in the promiscuous mode is employed. No routing algorithm is implemented: queries use a MAC-layer broadcast transmission, and information messages find their way back to the requesting node following a unicast path. Information messages exploit the request to send/clear to send (RTS/CTS) mechanism and MAC-level retransmissions, whereas query messages of broadcast nature do not use RTS/CTS and are never retransmitted. The channel operates at 11 Mb/s, and signal propagation is reproduced by a two-ray ground model. Simulations were run for 10 000 s.

In the aforementioned scenarios, our performance evaluation hinges upon the following quite-comprehensive set of metrics that are aimed at highlighting the benefits of using Hamlet in a distributed scenario:

- 1) The ratio between solved and generated queries, called solved-queries ratio;
- 2) The communication overhead;
- 3) The time needed to solve a query;
- 4) The cache occupancy.

V. Evaluation With Large-Sized Caches

Here, we evaluate the performance of Hamlet in a network of nodes with large storage capabilities, i.e., with caches that can store up to 50% of all information items. Because such characteristics are most likely found in vehicular communication devices, tablets, or smartphones, the network environments under study are the City and Mall scenarios. As discussed in Section IV, in this case, the Hamlet framework is employed to compute the caching

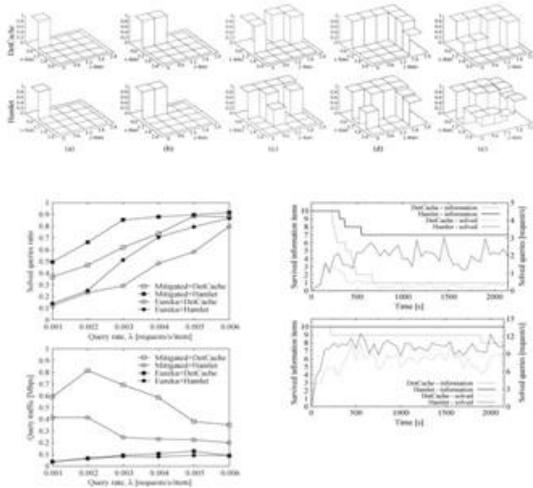
time for information chunks retrieved by nodes, with the goal of improving the content distribution in the network while keeping the resource consumption low. We first compare Hamlet's performance to the results obtained with a deterministic caching strategy, called DetCache, which simply drops cached chunks after a fixed amount of time.

Then, we demonstrate the effectiveness of Hamlet in the specific task of information survival. In all tests, we assume $I = 10$ items, each comprising $C = 30$ chunks. All items have identical popularity, i.e., all items are requested with the same rate $\lambda = \Lambda/I$ by all network nodes. The choice of equal request rates derives from the observation that, in the presence of nodes with a large-sized memory, caching an information item does not imply discarding another information item; thus, the caching dynamics of the different items are independent of each other and only depend on the absolute value of the query rate. It follows that considering a larger set of items would not change the results but only lead to more time-consuming simulations.

Each query includes 20 B plus 1 B for each chunk request, whereas information messages include a 20-B header and carry a 1024-B information chunk. The maximum caching time MC is set to 100 s, unless otherwise specified. Queries for chunks that are still missing are periodically issued every 5 s until either the information is fully retrieved or a timeout that is set to 25 s expires.

5.1 Benchmarking Hamlet

We set the deterministic caching time in DetCache to 40 s, and we couple DetCache and Hamlet with both the mitigated flooding and Eureka techniques for query propagation. We are interested in the following two fundamental metrics: 1) the ratio of queries that were successfully solved by the system and 2) the amount of query traffic that was generated. The latter metric, in particular, provides an indication of the system effectiveness in preserving locally rich information content: if queries hit upon the sought information in one or two hops, then the query traffic is obviously low. However, whether such a wealth of information is the result of a resource-inefficient cache-all-you-see strategy or a sensible cooperative strategy, e.g., the approach fostered by Hamlet, remains to be seen. Thus, additional metrics that are related to cache occupancy Fig: 5.1 Mall: Solved –queries ratio (top) and query traffic (bottom) with different schemes versus content request rate.



These values are very close to the DetCache caching time of 40 s, showing that Hamlet improves information survival by better distributing content in the network and not by simply caching them for longer periods of time.

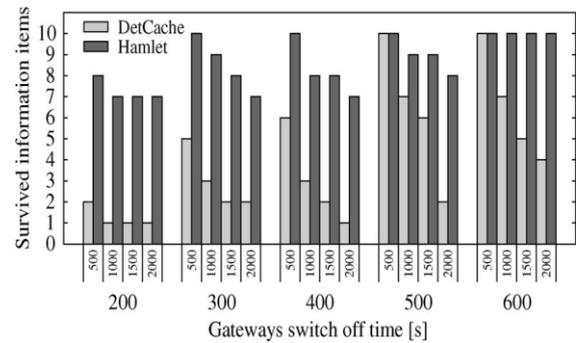


Fig 5.2 Information survival in the Mall (top) and

5.2 Information Survival

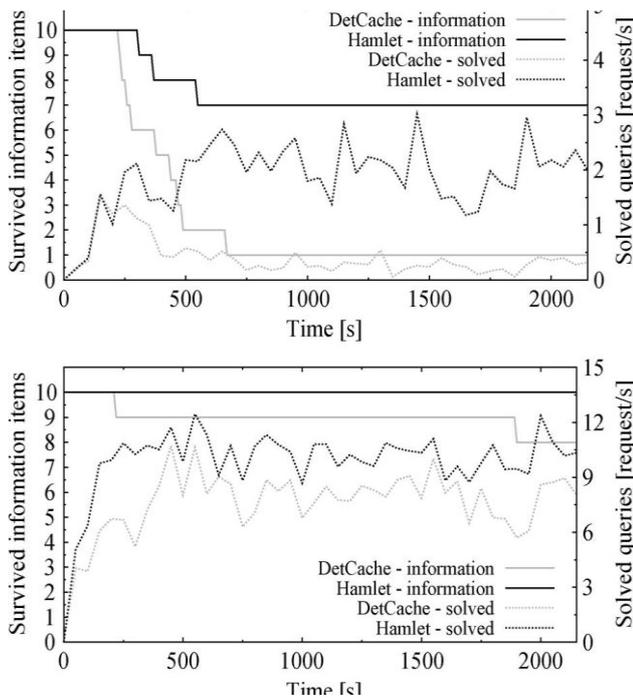


Fig5.3 Mall: Information survival for different

gateway switch-off times. The smaller numbers on the x-axis indicate the landmark time (in seconds) to which the number of survived items refers.(computed from the start of the simulation). Clearly, the later the gateways are shut down, the higher the probability of information survival, because the information has more time to spread through the network. We observe that Hamlet can maintain information presence equal to 100% if the information is given enough time to spread, i.e., gateways are disabled after 600 s or more, whereas DetCache loses half the items within the first 2000 s of simulation. We could wonder whether caching times give the edge to either Hamlet or DetCache. However, the average caching time in Hamlet ranges from 37 s to 45 s, depending on the gateway switch-off times and on the specific information item considered.

VI. Evaluation With Small-Sized Caches

We now evaluate the performance of Hamlet in a network where a node cache can accommodate only a small portion of the data that can be retrieved in the network. As an example, consider a network of low-cost robots that are equipped with sensor devices, where maps that represent the spatial and temporal behavior of different phenomena may be needed by the nodes and have to be cached in the network. We thus consider the memory-constrained scenario introduced in Section V and employ the Hamlet framework to define a cache replacement strategy. In such a scenario, the caching dynamics of the different information items become strongly intertwined. Indeed, caching an item often implies discarding different previously stored content, and as a consequence, the availability of one item in the proximity of a node may imply the absence of another item in the same area. Thus, in our evaluation, it is important to consider a large number of items, as well as to differentiate among these items in terms of popularity. We consider an overall pernode query rate $\Lambda = 0.1$ and sets of several hundreds of items. We assume that popularity levels q_i are distributed according to the Zipf law, which has been shown to fit popularity curves of content in different kinds of networks [29]. When not stated otherwise, the Zipf distribution exponent is set to 0.5. Such a value was selected, because it is close to the values observed in the real world [29], and the skewness that it introduces in the City (bottom) scenarios. The temporal behavior of the survived information and solved queries when the gateway nodes are switched off at $t = 200$ s.

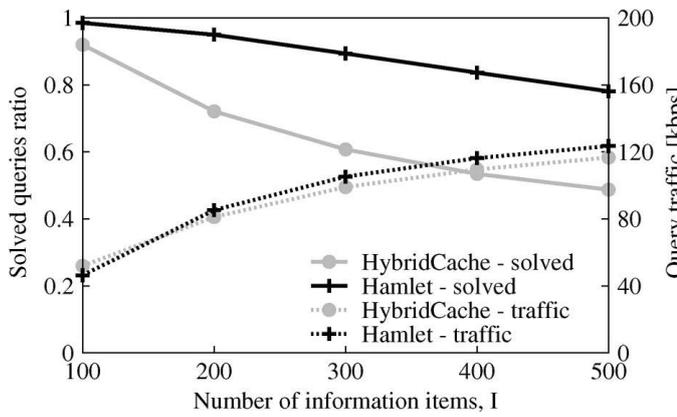


Fig 6.1 Static memory-constrained nodes: Solved-queries ratio and query traffic as the information set size varies, with Hybrid Cache and Hamlet.

popularity distribution is already sufficient to make differences emerge between the caching schemes that we study. In any case, we provide an analysis of the impact of the Zipf exponent at the end of this section. We assume that nodes can cache at most ten items, which correspond to a percentage between 2% and 10% of the entire information set, depending on the considered value of I. In addition, we set $C = 1$ to account for the smaller size of information items typically exchanged by memory-constrained nodes and MC to 300 s, because the increased network connectivity prolongs the reliability of information presence estimation.

6.1. Benchmarking Hamlet

Let us first focus on the memory-constrained scenario outlined in Section V with static nodes. Fig. 6.1 solved queries ratio and the overall query traffic versus the information set size. We observe that Hamlet reacts better to the growth of the number of items than HybridCache, without incurring any

penalty in terms of network load, as shown by the similar query traffic generated by the two schemes.

Observing the performance of Hamlet and HybridCache on a per-item basis allows a deeper understanding of the results.

In Fig. 6.2 (a), we show the solving ratio of the queries for each item when $I = 300$. Along the x-axis, items are ordered in decreasing order of popularity, with item 1 representing the most sought-after information and item 300 the least requested information. Unlike Hamlet, HybridCache yields extremely skewed query solving ratios for the different content; a similar observation also applies to the time needed to solve queries,

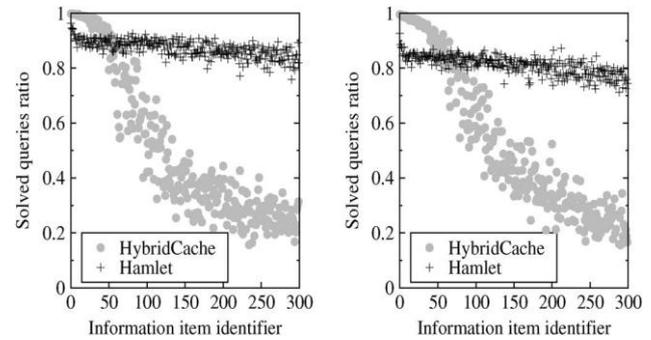


Fig. 6.3. Memory-constrained mobile nodes: Query-solving ratio for each information item when using HybridCache and Hamlet, with $I = 300$. The plots refer to v_m that is equal to 1 m/s (left) and 15 m/s (right).

We now compare the performance of HybridCache and Hamlet in the scenario with memory-constrained mobile nodes. We test the two schemes when $I = 300$ and for an average node speed v_m equal to 1 and 15 m/s.

The solved-queries ratio recorded with HybridCache and Hamlet on a per-item basis are shown in Fig. 13. Comparing the left and right plots, we note that the node mobility, even at high speed, does not seem to significantly affect the results due to the high network connectivity level. The spatial redistribution of content induced by node movements negatively affects the accuracy of Hamlet's estimation process, which explains the slight reduction in the solved query ratio at 15 m/s. That same movement favors HybridCache, at least at low speed, because it allows unpopular information to reach areas that are far from the gateway. However, the difference between the two schemes is evident, with Hamlet solving an average of 20% requests more than HybridCache, when nodes move at 15 m/s. even if it represents two thirds of the whole information set.

Instead, Hamlet achieves, in a completely distributed manner, a balanced networkwide utilization of node caches. Indeed, the results of Hamlet are very close to

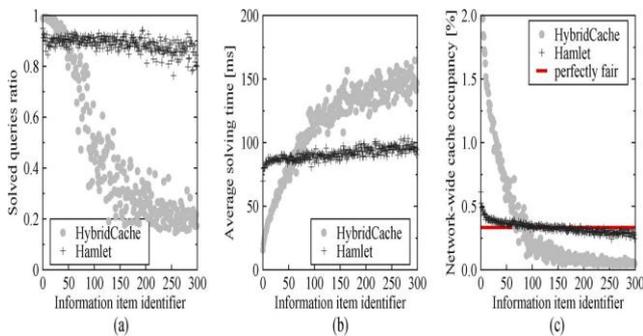


Fig. 6.2. Static memory-constrained nodes. (a) Query-solving ratio, (b) time, and (c) average networkwide cache occupancy for each item when using HybridCache and Hamlet, with $I = 300$. In (c), the red horizontal line represents perfect fairness in cache occupancy among different items.

the most even cache occupancy that we can have, represented by the horizontal red line in the plot and corresponding to the case where the total network storage capacity is equally shared among the I items

VII. Results



Fig 1. Out Put Screen for Data Caching in Mobile nodes

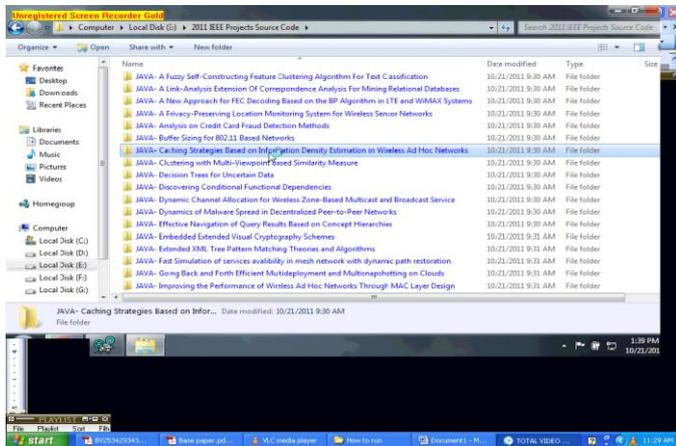


Fig 2. Screen For Using Java Technology



Fig 3. Mobile nodes in wireless adhoc-networks

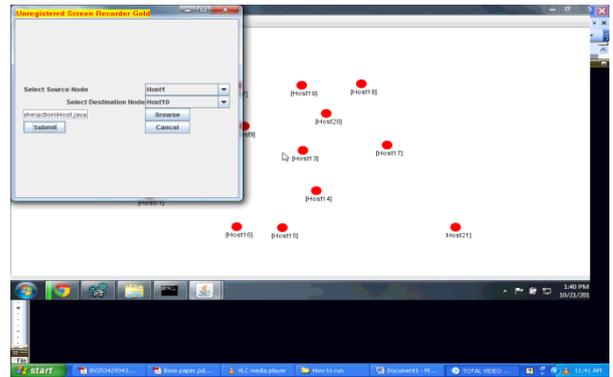


Fig 4. Screen from Source Host to Destination Host

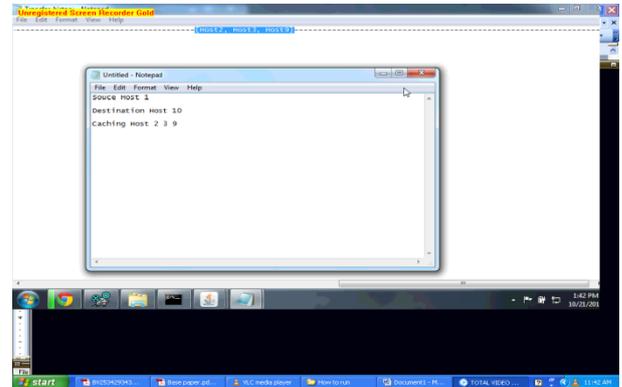


Fig 5. Screen for Data Caching between Hosts 2, 3, 9

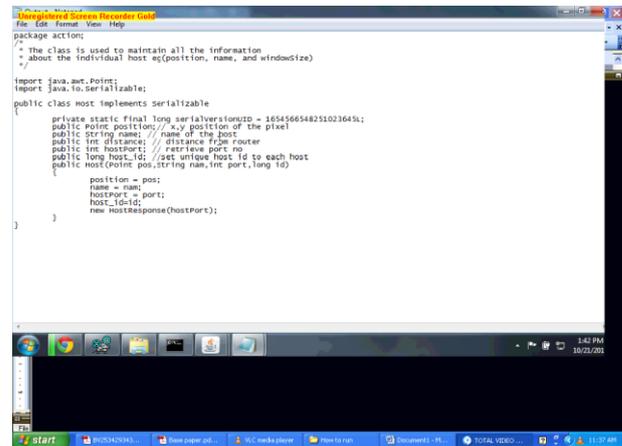


Fig 6. Using Java package function for each caching node position

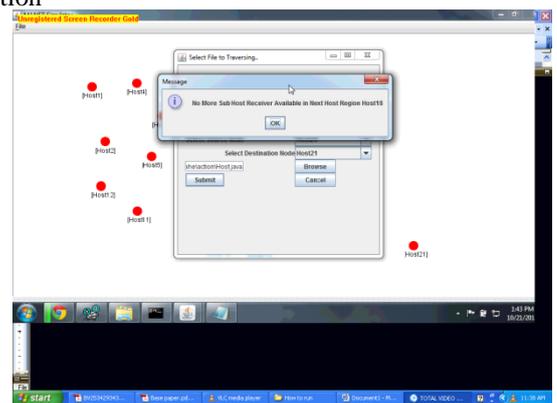


Fig 7. Screen for No More Sub Hosts for caching the data

VIII. Conclusion

We have introduced Hamlet, which is a caching strategy for ad hoc networks whose nodes exchange information items in a peer-to-peer fashion. Hamlet is a fully distributed scheme where each node, upon receiving a requested information, determines the cache drop time of the information or which content to replace to make room for the newly arrived information. These decisions are made depending on the perceived "presence" of the content in the node's proximity, whose estimation does not cause any additional overhead to the information sharing system. Data caching strategy for ad hoc networks whose nodes exchange information items in a peer-to-peer fashion. Data caching is a fully distributed scheme where each node, upon receiving requested information, determines the cache drop time of the information or which content to replace for the newly arrived information. We have developed a paradigm of data caching techniques to support effective data access in ad hoc networks. In particular, we have considered memory capacity constraint of the network nodes. We have developed efficient data caching algorithms to determine near optimal cache placements to maximize reduction in overall access cost. Reduction in access cost leads to communication cost savings and hence, better bandwidth usage and energy savings.

However, our simulations over a wide range of network and application parameters show that the performance of the caching algorithms. Presents a distributed implementation based on an approximation algorithm for the problem of cache placement of multiple data items under memory constraint. The result is the creation of content diversity within the nodes neighborhood so that a requesting user likely finds the desired information nearby. We simulate our caching algorithms in different ad hoc network scenarios and compare them with other caching schemes, showing that our solution succeeds in creating the desired content diversity, thus leading to a resource-efficient information access. We showed that, due to Hamlet's caching of information that is not held by nearby nodes, the solving probability of information queries is increased, the overhead traffic is reduced with respect to benchmark caching strategies, and this result is consistent in vehicular, pedestrian, and memory-constrained scenarios. Conceivably, this paper can be extended in the future by addressing content replication and consistency.

The procedure for information presence estimation that was developed in Hamlet can be used to select which content should be replicated and at which node (even if such a node did not request the content in the first place). In

addition, Hamlet can be coupled with solutions that can maintain consistency among copies of the same information item cached at different network.

References

- [1] J. Wortham (2009, Sep.). Customers Angered as iPhones Overload AT&T. The New York Times. [Online]. Available: <http://www.nytimes.com/2009/09/03/technology/companies/03att.html>
- [2] A. Lindgren and P. Hui, "The quest for a killer app for opportunistic and delay-tolerant networks," in Proc. ACM CHANTS, 2009, pp. 59–66.
- [3] P. Padmanabhan, L. Gruenwald, A. Vallur, and M. Atiquzzaman, "A survey of data replication techniques for mobile ad hoc network databases," VLDB J., vol. 17, no. 5, pp. 1143–1164, Aug. 2008.
- [4] A. Derhab and N. Badache, "Data replication protocols for mobile ad hoc networks: A survey and taxonomy," IEEE Commun. Surveys Tuts., vol. 11 no. 2, pp. 33–51, Second Quarter, 2009.
- [5] B.-J. Ko and D. Rubenstein, "Distributed self-stabilizing placement of replicated resources in emerging networks," IEEE/ACM Trans. Netw., vol. 13, no. 3, pp. 476–487, Jun. 2005.
- [6] G. Cao, L. Yin, and C. R. Das, "Cooperative cache-based access in ad hoc networks," Computer, vol. 37, no. 2, pp. 32–39, Feb. 2004.
- [7] C.-Y. Chow, H. V. Leong, and A. T. S. Chan, "GroCoca: Group-based peer-to-peer cooperative caching in mobile environment," IEEE J. Sel. Areas Commun., vol. 25, no. 1, pp. 179–191, Jan. 2007.
- [8] T. Hara, "Cooperative caching by mobile clients in push-based information systems," in Proc. CIKM, 2002, pp. 186–193.
- [9] L. Yin and G. Cao, "Supporting cooperative caching in ad hoc networks," IEEE Trans. Mobile Comput., vol. 5, no. 1, pp. 77–89, Jan. 2006.
- [10] N. Dimokas, D. Katsaros, and Y. Manolopoulos, "Cooperative caching in wireless multimedia sensor networks," ACM Mobile Netw. Appl., vol. 13, no. 3/4, pp. 337–356, Aug. 2008.
- [11] Y. Du, S. K. S. Gupta, and G. Varsamopoulos, "Improving on-demand data access efficiency in MANETs with cooperative caching," Ad Hoc Netw., vol. 7, no. 3, pp. 579–598, May 2009.
- [12] Y. Zhang, J. Zhao, and G. Cao, "Roadcast: A popularity-aware content sharing scheme in

VANETs,” in Proc. IEEE Int. Conf. Distrib. Comput. Syst., Los Alamitos, CA, 2009, pp. 223–230.

- [13] W. Li, E. Chan, and D. Chen, “Energy-efficient cache replacement policies for cooperative caching in mobile ad hoc network,” in Proc. IEEE WCNC, Kowloon, Hong Kong, Mar. 2007, pp. 3347–3352.

AUTHORS LIST



Tulasi Rami Reddy Yeddula (M.Tech)

Received his Master’s Degree in CSE in MKU University, Madurai and Pursuing Masters of Technology in Computer Science & Engineering in Kottam College of Engineering, Kurnool, affiliated to JNTU Anantapur. His research areas of interest are Computer Networks and Data Mining.



K.GOPINATH M.Tech (Ph.D), Completed

his B.Tech ,Computer Science in (JNTU, Anantapur) 1999 and M.tech, Computer Science (JNTU, Hyd) 2002 .Presently Pursuing Ph.D in Data Mining JNTUniversity, Hyderabad . His area of intrest are Computer Networks and Data Mining.

Con S – \mathcal{K} – EP Matrices and Its Weighted Generalized Inverse

S.Krishnamoorthy¹ and B.K.N.Muthugoba²

*Research Scholar Ramanujan Research Centre, Department of Mathematics,
 Govt. Arts College (Autonomous), Kumbakonam, Tamilnadu, India.*

ABSTRACT: If A is a con s – \mathcal{K} – EP matrix, then the weighted generalized inverse of A (with respect to the given matrices M, N) is a matrix which satisfies $AA^\dagger A = A$, $A^\dagger A A^\dagger = A^\dagger$ and that MAA^\dagger and that $AA^\dagger N$ are symmetric under certain conditions on M, N .

It is shown that the weighted generalized inverse exists if and only if $AN A^T M A = A$, in which case the inverse is $N^T A^T M^T$. When M, N are identity matrices, this reduces to the well known result that the weighted generalized inverse of a con-s-k-EP matrix when it exists, must be A^T .

Keywords: con-s-k-EP matrix, generalized inverse, weighted generalized inverse.

AMS classification: 15A09, 15A15, 15A57

I. Introduction

Let $C_{n \times n}$ be the space of $n \times n$ complex matrices of order n . Let C_n be the space of all complex n tuples. For $A \in C_{n \times n}$ Let $\bar{A}, A^T, A^*, A^S, \bar{A}^S, A^\dagger, R(A), N(A)$ and $\rho(A)$ denote the conjugate, transpose, conjugate transpose, secondary transpose, conjugate secondary transpose, Moore Penrose inverse range space, null space and rank of A respectively. A solution X of the equation $AXA = A$ is called generalized inverse of A and is denoted by A^- . If $A \in C_{n \times n}$ then the unique solution of the equations $A X A = A, X A X = X [AX]^* = AX, [XA]^* = XA[2]$ is called the Moore-Penrose inverse of A and is denoted by A^\dagger . A matrix A is called con s – \mathcal{K} – EP_r if $\rho(A) = r$ and $N(A) = N(A^T V K)$ (or) $R(A) = R(K V A^T)$. Throughout this paper let " \mathcal{K} " be the fixed product of disjoint transposition in $S_n = \{1, 2, \dots, n\}$ and k be the associated permutation matrix.

Let us define the function $\mathcal{K}(x) = (x_{k(1)}, x_{k(2)}, \dots, x_{k(n)})$

A matrix $A = (a_{ij}) \in C_{n \times n}$ is s - k - symmetric if

$a_{ij} = a_{n-k(j)+1, n-k(i)+1}$ for $i, j = 1, 2, \dots, n$. A matrix $A \in C_{n \times n}$ is said to be

Con –s-k- EP if it satisfies the condition

$A_x = 0 \iff A^S \mathcal{K}(x) = 0$ or equivalently $N(A) = N(A^T V K)$. In addition to that A is con-s-k-EP $\iff KVA$ is con-EP or AVK is con-EP and A is con-s-k-EP $\iff A^T$ is con-s-k-EP_r moreover A is said to be con-s-k-EP_r if A is con-s-k-EP and of rank r . For further properties of con-s-k-EP matrices one may refer [1].

Definition 1.1

Let $A, M, N \in C_{n \times n}$. The weighted Moore – Penrose inverse of A (with respect to M, N denoted by $A_{M, N}^\dagger$) is defined to be $n \times n$ matrix A^\dagger satisfying

- (i) $AA^\dagger A = A$
- (ii) $A^\dagger A A^\dagger = A^\dagger$
- (iii) $(MAA^\dagger)^T = MAA^\dagger$
- (iv) $(A^\dagger AN)^T = A^\dagger AN$

In case M,N are identity matrices then the matrix A^\dagger satisfying (i) – (iv) is simply the Moore Penrose inverse (denoted by A^\dagger) of A

Theorem 1.2

Let A be an nxn con-s-k-EP matrix. Then the following assertions are equivalent:

- (i) The Moore - Penrose inverse of A exists
- (ii) The Moore - Penrose inverse of A exists and equals A^T .
- (iii) $AA^T A = A$
- (iv) $AA^T A \leq A$
- (v) Any two rows of A are either identical or disjoint (ie, there is no column with a in both the rows)
- (vi) Any two columns of are either identical or disjoint.
- (vii) The number of ones in any 2x2 sub matrix of A is not 3.
- (viii) Any 2x2 submatrix of A admits a Moore Penrose inverse.
- (ix) There exist permutation matrix P,Q such that

$$PAQ = \begin{bmatrix} J_1 & 0 & \dots & 0 & 0 \\ 0 & J_2 & \dots & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \dots & J_t & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Where $J_1 \dots\dots\dots J_t$ are matrices (not necessarily square) of all ones.

- (x) There exists permutation matrices P, Q Such that

$$PAQ = \begin{bmatrix} I & C \\ D & DC \end{bmatrix}$$

Where C, D satisfy $CC^T \leq I, D^T D \leq I$.

- (xi) The main purpose of the present paper is to generalize some aspects of theorem 1.2 to the weighted case. The proof techniques are new and may be used to obtain results for matrices over more general structures. Thus most of our statements are valid for matrices over a distributive lattice. Whereas some require the structure of a completely ordered set. Such generalizations will be clear from the proofs. However, we have chosen to present the results only in the setting of con-s-k-EP matrices. In the next section we consider the question of the existence of a weighted Moore – Penrose inverse and give a formula for it when it exists

II. The Main Result

we begin by showing that under some conditions on M,N the inverse $A_{M,N}^\dagger$. when it exists is unique , we denote the row space of the matrix A by $R(A)$, and the column space by $R(A)$.

THEOREM 2.1

Let A be a con-s-k-EP matrix of order nxn and M,N be matrices order nxn. $R(KVA^T) = R(MA), R(KVA^T) = R(AN)$. There exist matrices x,y suchthat $XMA = A, ANY = A$. Then

- (a) $AN^T A^T = A NA^T, A^T M^T A = A^T MA$
- (b) $A^\dagger_{M,N}$ is unique.

Proof :

- (a) Let $A^\dagger = A^\dagger_{M,N}$ exists then

$$\begin{aligned} AN^T A^T &= AN^T A^T (A^T)^\dagger A^T \quad (\text{since } AA^\dagger A = A^T) \\ &= AA^\dagger AN A^T \quad (\text{since } A^\dagger AN \text{ is symmetric}) \\ &= AN A^T \quad (\text{since } AA^\dagger A = A) \end{aligned}$$

The proof of the remaining part of

(a) Is similar to the above

(b) Let, if possible A_1^\dagger, A_2^\dagger be two candidates for $A_{M,N}^\dagger$ then

$$\begin{aligned} A_1^\dagger AN &= A_1^\dagger AA_2^\dagger AN \quad (\text{since } A = AA_2^\dagger A) \\ &= A_1^\dagger AN A^T A_2^{\dagger T} \quad (\text{since } A_2^\dagger AN \text{ is symmetric}) \\ &= A_1^\dagger AN^T A^T A_2^{\dagger T} \quad (\text{using (a)}) \\ &= N^T A^T A_1^{\dagger T} A^T A_2^{\dagger T} \quad (\text{since } A_1^\dagger AN \text{ is symmetric}) \\ &= N^T A^T A_2^{\dagger T} \quad (\text{since } AA_1^\dagger A = A) \\ &= A_2^\dagger AN \quad (\text{since } A_2^\dagger AN \text{ is symmetric}). \end{aligned}$$

Thus $A_1^\dagger ANY = A_2^\dagger AN Y$ and hence $A_1^\dagger A = A_2^\dagger A$ (since $ANY = A$). it follows

$$\begin{aligned} A_1^\dagger A A_1^\dagger &= A_2^\dagger A A_1^\dagger \quad \text{and therefore} \\ A_1^\dagger &= A_2^\dagger A A_1^\dagger \quad (1) \end{aligned}$$

Now

$$\begin{aligned} MA A_1^\dagger &= M A A_2^\dagger A A_1^\dagger \quad (\text{since } A = AA_2^\dagger A) \\ &= A_2^{\dagger T} A^T M^T A A_1^\dagger \quad (\text{since } MA A_2^\dagger \text{ is symmetric}) \\ &= A_2^{\dagger T} A^T M A A_1^\dagger \quad (\text{using (a)}) \\ &= A_2^{\dagger T} A^T A_1^{\dagger T} A^T M^T \quad (\text{since } MA A_1^{\dagger T} \text{ is symmetric}) \\ &= A_2^{\dagger T} A^T M^T \quad (\text{since } AA_1^\dagger A = A) \\ &= M A A_2^\dagger \quad (\text{since } MA A_2^\dagger \text{ is symmetric}) \end{aligned}$$

It follows that $X M A A_1^\dagger = X M A A_2^\dagger$ and hence

$$A A_1^\dagger = A A_2^\dagger \quad (\text{since } XMA = A)$$

Therefore $A_2^\dagger A A_1^\dagger = A_2^\dagger A A_2^\dagger$ and thus

$$A_2^\dagger A A_1^\dagger = A_2^\dagger$$

It follows from (1) (2) that $A_1^\dagger = A_2^\dagger$ and the proof is complete.

Example

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad A_1^\dagger = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

Take $M=I$ and N to be the 2×2 zero matrix. Then it can be verified that both A_1^\dagger, A_2^\dagger satisfy all conditions in definition 1 and therefore the weighted generalized inverse is not unique in this example observe that here the condition of theorem 2.1 is not satisfied

The next result will be used in the sequence

Lemma 1:

Let A be an $n \times n$ matrix. then $A \leq KVA A^T AVK$

Proof :

Let $B = KVA^T AVK$ we must show that $a_{ij} \leq b_{ij}$ for all i, j this is $b_{ij} = KV [\sum_{k=1}^n \sum_{i=j}^m a_{ik} a_{lk} a_{lj}]$
 VK (3)

If we set $k=j, l=i$ then $a_{ik} a_{lj} a_{li} = a_{ij}^3 = a_{ij}$. It follows from (3) that $b_{ij} \geq a_{ij}$ and the proof complete .

The following is the main result of this section.

Theorem 2.2

(a) $R(KVA^T) = R(MA), R(KVA^T) = R(AN)$

(b) $M \geq I, N \geq I$

Then the following assertions are equivalent:

(i) $A_{M,N}^\dagger$ exists

(ii) Any one of the following holds

(1) $ANA^T MA = A$ (2) $AN^T A^T MA = A$

(3) $ANA^T M^T A = A$ (4) $AN^T A^T M^T A = A$

And thus $A_{M,N}^\dagger = N^T A^T M^T$

(iii) Any two rows of A are either identical or disjoint, and $ANA^T = AA^T$

(iv) Any two columns of A are either identical or disjoint and $ANA^T = AA^T$

$A^T MA = A^T A.$

Proof:

(i) \Rightarrow (ii)

Suppose $A^\dagger = A_{M,N}^\dagger$ exists since the number of con-s- k -EP matrices of a given order is finite there exists integers $k \geq 1, s \geq 1$ such that

$(ANA^T M^T)^k = (ANA^T M^T)^{k+s}$ (4)

Without loss we can assume $s > 1$ for if $s=1$ then (4) clearly holds for $S=2$ as well as now left multiplying equation (4) by A^\dagger and then using the fact $A^\dagger AN$ is symmetric, we get

$N^T A^T A^\dagger M^T (ANA^T M^T)^{k-1} = N^T A^T A^\dagger M^T (ANA^T M^T)^{k-1+s}$

Left multiply the above equation by Y^T and then use $AGA = A, ANY = A$ to get $A^T M^T (ANA^T M^T)^{k-1} = A^T M^T (ANA^T M^T)^{k-1+s}$.

Left multiply the above equation by A^\dagger and then use the facts that MAA^\dagger is symmetric and the $AA^\dagger A = A$ to get

$M(ANA^T M^T)^{k-1} = M(ANA^T M^T)^{k-1+s}$. Finally left multiply the above equation by X and use $XMA = A$ to get

$(ANA^T M^T)^{k-1} = (ANA^T M^T)^{k-1+s}$. Continuing this way we may assume $k=1$, without loss of generality and therefore

$ANA^T M^T = (ANA^T M^T)^{s+1}$

Starting with the above equation, we get the following chain of implications:

$\Rightarrow A^\dagger ANA^T M^T = A^\dagger ANA^T M^T (ANA^T M^T)^s$

$\Rightarrow N^T A^T A^\dagger M^T = N^T A^T A^\dagger M^T (ANA^T M^T)^s$

$\Rightarrow N^T A^T M^T = N^T A^T M^T (ANA^T M^T)^s$

$\Rightarrow Y^T N^T A^T M^T = Y^T N^T A^T M^T (ANA^T M^T)^s$

$\Rightarrow A^T M^T = A^T M^T (ANA^T M^T)^s$

$$\Rightarrow A^T M^T X^T = A^T M^T (ANA^T M^T)^S X^T$$

$$\Rightarrow A^T = A^T M^T (ANA^T M^T)^{S-1} ANA^T$$

And therefore

$$A = (ANA^T M)A \quad (5)$$

By Lemma 1.

$A \leq A A^T A$, and so since $M \geq I$, $N \geq I$ we have

$$A \leq AN^T A^T MA \quad (6)$$

And hence post multiplying by $N^T A^T MA$ we get

$$AN^T A^T MA \leq (AN^T A^T M)^2 A \quad (7)$$

Repeated post multiplication of (7) by $N^T A^T MA$ gives

$$\begin{aligned} A &\leq AN^T A^T MA \leq (AN^T A^T M)^2 A \\ &\leq (AN^T A^T M)^3 A \leq \dots \leq (AN^T A^T M)^2 A = A \end{aligned}$$

Where the equality follows in view of (5)

$$A = AN^T A^T MA = ANA^T MA = ANA^T M^T A = AN^T A^T M^T A, \quad (8)$$

Where the last three equalities follow in view of Theorem 2.1(a) now we show $A^\dagger = A_{M,N}^\dagger = N^T A^T M^T$ We have shown

$$AA^\dagger A = A, \text{ but } A^\dagger AA^\dagger = N^T A^T M^T AN^T A^T M^T = N^T A^T M^T \text{ (in view of (8)) and}$$

$$MAG = MAN^T A^T M^T$$

$$= MANAM^T \text{ (Using Theorem 2.1 (a))}$$

$$= (MAN^T A^T M^T)^T = (MAA^\dagger)^T$$

So MAA^\dagger is symmetric showing $A^\dagger AN$ symmetric is similar .

$$\text{so } A_{M,N}^\dagger = N^T A^T M^T$$

(ii) \Rightarrow (i) let $ANA^T MA = A$. By Lemma 1

$A \leq AA^T A$ as $M \geq I$, $N \geq I$ we have

$A \leq AA^T A \leq ANA^T A \leq ANA^T MA = A$ thus

$$= AA^T A = A = ANA^T A \quad (9)$$

The second part of the above equation gives

$$AA^T = ANA^T AA^T$$

$$= ANA^T \text{ (using the first part of (9))}$$

$$= AN^T A^T \text{ (since } AA^T \text{ is symmetric)}$$

Similarly it can be shown that $A^T MA = A^T M^T A = A^T A$

Now using these facts equation (9) and the assumption one can easily see that $A_{M,N}^\dagger = N^T A^T M^T$.

(ii) \Rightarrow (iii) without loss we take $ANA^T MA = A$

Suppose two rows of A, say the i th and j th are not disjoint. Then there exists k such that $a_{ik} = a_{jk} = 1$. now if

$a_{ir} = 1$ for some r , then we have

$$a_{jr} \geq a_{jk} n_{kk} a_{ik} M_{ir} a_{ir} = 1$$

and hence $a_{jr} = 1$. thus the i th row of A is entry wise dominated by the j th row. Similarly we can shown that the j th row of A is entry wise dominated by the i th row and hence the two rows must be identical.

The proof of the remaining part is essentially contained in the proof of (ii) \Rightarrow (i)

(iii) \Rightarrow (ii) : let $B = AA^T A$ and suppose $b_{ij} = 1$. So there exists $l_1 l_2$ such that

$$a_i b_1 = a l_2 t_1 = a l_2 j = 1$$

Now observe that the l_1 the column of A is nonzero . so by hypothesis we have that the i th row of A is equal to the l_2 the row of A. but we also have $a_{i_2 j} = 1$.

So , $a_{ij}=1$ and therefore $AA^T A \leq A$. it follows by lemma 1 that $A = AA^T A$.

It follows by lemma 1 that $A = AA^T A$. Since $ANA^T = AA^T$, $A^T MA = A^T A$.

Then $ANA^T MA = AA^T MA = AA^T A = A$

and (ii) is proved. The equivalence of (iv) and (ii) is proved similarly that completes the proof of the theorem.

An examination of the proof of theorem 2.2 reveals that condition (b) may be replaced by the weaker condition $ANA^T MA \geq A$.

We now provide a proof of theorem 1.1

Proof of the 1.1 . the equivalence of (i) – (iv) of the theorem essentially follows from theorem 2.2 by setting $M=N=I$. the equivalence of (v) and (vii) is easy to prove and so is the equivalence of (vii) and (viii). The implication (v) \Rightarrow (ix) and (ix) \Rightarrow (iii) are easy to prove thus we have shown that assertions (i) –(ix) are equivalent.

It is easy to see that assertions (i) –(ix) are equivalent.

It is easy to see that (ix) \Rightarrow (x). if (x) holds then it can be verified that A^T is the Moore Penrose inverses of A and this (i) holds that completes the proof

We remark that all the assertions in theorem 1.1 except (vii) .(viii) are essentially contained in the literatures. See [2,4,5] however, we have given proofs for completes the proof.

As shown in theorem 1.1 if A admits a Moore Penrose inverse, then it must be A^T some times it happens that the weighted Moore Penrose inverse $A_{M,N}^\dagger = A^T$, the trivial case being $M = N = I$. So the obvious question is whether we can precisely point out the cases when $A_{M,N}^\dagger = A^T$. To answer this question we need the following result.

Theorem 2.3:

Let A,M,N be as in Theorem 2.2 then the following are equivalent

- (i) $A_{M,N}^\dagger$ exists
- (ii) There exists permutation matrices P and A such that

$$\bar{A} = PAQ = \begin{pmatrix} J_1 & 0 & \dots & 0 & 0 \\ 0 & J_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & & J_K & 0 \\ 0 & 0 & & 0 & 0 \end{pmatrix}$$

$$\bar{N} = \begin{pmatrix} N_{ij} & \dots & N_{I1,K+1} \\ \vdots & & \vdots \\ N_{K+1} & \dots & N_{K+1,K+1} \end{pmatrix}$$

A conformed partition and then using (10) we see that all the blocks $N_{ij}=0$

$I \leq i, j \leq k, i \neq j$ and that $N_{ij} \neq 0, 1 \leq i \leq k$. We have a similar conclusion regarding \bar{M} .

It is easy to see that $\bar{A}^\dagger = \bar{N}^T \bar{A}^T \bar{M}^T$ is the weighted Moore – Penrose inverse of \bar{A} with respect to \bar{M}, \bar{N} . But

$$\begin{aligned} \bar{A}^\dagger &= (Q^T N^T Q) (Q^T A^T P^T) (P M^T P^T) = Q^T (N^T A^T M^T) P^T \\ &= Q^T A^\dagger P^T \end{aligned}$$

Now carrying out the block multiplication in the equation $\bar{A}^\dagger = \bar{N}^T \bar{A}^T \bar{M}^T$. We see that \bar{A}^\dagger is of the form gives in the statement

Now by (ii), the proof of (i) \Rightarrow (ii) is complete.

Conversely, suppose (ii) holds defining \bar{A}^\dagger as in the statement of the theorem, it is easy to check that $\bar{A}\bar{A}^\dagger\bar{A} = \bar{A}$.

Since $\bar{A}^\dagger = \bar{N}^T \bar{A}^T \bar{M}^T$.

We have $\bar{A}\bar{N}^T \bar{A}^T \bar{M}^T \bar{A} = \bar{A}$.

This implies $PA\bar{N}^T \bar{A}^T \bar{M}^T AQ = PAQ$.

Therefore $AN^T A^T M^T A = A$ and thus $A_{M,N}^\dagger$ by theorem 2.2.

As a simple corollary we state the following result without proof.

Corollary 1:

Let A,M,N be as in theorem 2.2 then $A_{M,N}^\dagger = A^T$ if and only if condition (ii) of theorem 2.3 is satisfied with the additional proviso that \bar{M} and \bar{N} are block diagonal .

We also have the following

Corollary 2:

Let A,M,N be as in Theorem 2.2 and further suppose A has no zero row or zero column then if $A_{M,N}^\dagger$ exists it equals A^T .

Proof :

Observe that \bar{A} has no diagonal zero block. Hence by Theorem 2.3 \bar{M}, \bar{N} are block diagonal furthermore that $A_{M,N}^\dagger$

Exists implies that \bar{A}^\dagger exists . the result now follows by corollary 1. we conclude with an example which shows that the condition that A has no zero row or column is necessary in corollary 2.

Example:

Let $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ $N = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ $M = I$

Then $ANA^T MA = A$ but $NA^T M = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \neq A^T$

We sincerely thank the referee for a careful reading of the manuscript and for suggesting theorem 2.3 .

References

1. S. krishnamoorthy, k. gunasekaran and B.K.N muthugobal , con-s-k-EP matrices , journal of mathematical sciences and engineering applications vol .5, no.1, 211, pp.353 – 364.
2. C.R .RAO and S.K.Mitra “ generalized inverse of matrices and its applications” wiley and sons new York 1971.
3. A.B.Isral and T.N.E.Greviue , “ generalized inverses theory and applications” wiley and sons, newyork 1974.

Bio Diesel from Waste Vegetable Oil

Dr.B.Karunanithi¹, K.Bogeshwaran², Manasa Tripuraneni³

¹Professor, Department of Chemical Engineering, SRM University, Kattankulathur,603203,India.

²Assistant Professor, Department of Chemical Engineering, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, 600062,India.

³Assistant Professor, Department of Mechanical Engineering, Eswar College of Engineering, Guntur, 522601, India.

Abstract: The world is confronted with the twin crises of fossil fuel depletion and environmental degradation. The indiscriminate extraction and consumption of fossil fuels have led to a reduction in petroleum reserves. Alternative fuels, energy conservation and management, energy efficiency and environmental protection have become important in recent years. The increasing import bill has necessitated the search for liquid fuels as an alternative to diesel, which is being used in large quantities in transport, agriculture, industrial, commercial and domestic sectors. Biodiesel obtained from vegetable oils has been considered a promising option. In this paper, an attempt has been made to produce biodiesel from waste vegetable oils and the properties of the produced biodiesel have been studied including its emission characteristics. A four stroke, single cylinder is used to study the emission and performance characteristics. The large scale production of Biodiesel and its economic aspects have also been discussed in brief.

Keywords: Alternative fuels, Bio diesel, Emission Characteristics, Environmental Protection, Four stroke engines, Vegetable oil,

I. INTRODUCTION

1.1 Introduction:

In recent times, the world has been confronted with an energy crisis due to depletion of resources and increased environmental problems. The situation has led to the search for an alternative fuel, which should be not only sustainable but also environment friendly. For developing countries, fuels of bio-origin, such as alcohol, vegetable oils, biomass, biogas, synthetic fuels, etc. are becoming important. Such fuels can be used directly, while others need some sort of modification before they are used as substitute of conventional fuels. As per an estimate, India consumed about 40.34 million tons of diesel in 2000–2001, which was 43.2% of the total consumption of petroleum products [1], and two-thirds of the demand was met by import costing about Rs.200 billion. With an expected growth rate of diesel consumption of more than 14% per annum, shrinking crude oil reserves and limited refining capacity, India will have to depend heavily on imports of crude. From the point of view of protecting the global environment and the concern for long term supplies of conventional diesel fuels, it becomes necessary to develop alternative fuels comparable with conventional fuels.

Diesel fuel is largely utilized in the transport, agriculture, commercial, domestic, and industrial sectors for the generation of power/mechanical energy, and the substitution of even a small fraction of total consumption by alternative fuels will have a significant impact on the economy and the environment. Of the alternative fuels, biodiesel obtained from vegetable oils holds good promises as an eco-friendly alternative to diesel fuel.

1.2 Alternate Diesel Fuels:

Alternative fuels should be easily available, environment friendly and techno-economically competitive. One of such fuels is triglycerides (vegetable oils/animal fats) and their derivatives. Vegetable oils, being renewable, are widely available from a variety of sources and have low sulphur contents close to zero, and hence cause less environmental damage (lower greenhouse effect) than diesel. Besides, vegetable oils and their derivatives are produced widely in the country for food and other purposes.

1.3 Triglycerides as diesel fuels

The use of vegetable oils, such as palm, Soya bean, sunflower, peanut, and olive oil, as alternative fuels for diesel engines dates back almost nine decades, but due to the rapid decline in crude oil reserves, it is again being promoted in many countries. Depending upon the climate and soil conditions, different countries are looking for different types of vegetable oils as substitutes for diesel fuels. For example, Soya bean oil in the US, rapeseed and sunflower oils in Europe, palm oil in South-east Asia (mainly Malaysia and Indonesia) and coconut oil in the Philippines are being considered. The production of oil seeds, percentage oil recovery and their respective cost are given in [2], which indicates that the use of vegetable oils as sources of diesel would require more efforts to increase the production of oil seeds and to develop new and more productive plant species with high yield of oil. Besides, some species of plants yielding non-edible oils, e.g. jatropha, karanja and pongamia may play a significant role in providing resources. Both these plants may be grown on a massive scale on agricultural/degraded/waste lands, so that the chief resource may be available to produce biodiesel on 'farm scale'.

1.4 Properties of Vegetable Oils as Fuel

The fuel properties of vegetable oils [2,4] indicate that the kinematics viscosity of vegetable oils varies in the range of 30–40 cSt at 38 8C. The high viscosity of these oils is due to their large molecular mass in the range of 600–900, which is about 20 times higher than that of diesel fuel. The flash point of vegetable oils is very high (above 200. 8C). The volumetric heating values are in the range of 39–40 MJ/kg, as compared to diesel fuels (about 45 MJ/kg). The presence of chemically bound oxygen in vegetable oils lowers their heating values by about 10%. The cetane numbers are in the range of 32–40.

1.5 Utilization of vegetable oils as fuels

It has been found that these vegetable oils can be used as diesel fuels in conventional diesel engines, but this leads to a number of problems related to the type and grade of oil and local climatic conditions. The injection, atomization and combustion characteristics of vegetable oils

in diesel engines are significantly different from those of diesel. The high viscosity of vegetable oils interferes with the injection process and leads to poor fuel atomization. The inefficient mixing of oil with air contributes to incomplete combustion, leading to heavy smoke emission, and the high flash point attributes to lower volatility characteristics. These disadvantages, coupled with the reactivity of unsaturated vegetable oils, do not allow the engine to operate trouble free for longer period of time. These problems can be solved, if the vegetable oils are chemically modified to biodiesel, which is similar in characteristics to diesel.

1.6 Fuel properties of biodiesel

The properties of biodiesel and diesel fuels, in general, show many similarities, and therefore, biodiesel is rated as a strong candidate as an alternative to diesel. This is due to the fact that the conversion of triglycerides into methyl or ethyl esters through the transesterification process reduces the molecular weight to one-third, reduces the viscosity by about one-eighth, and increases the volatility marginally. Biodiesel contains 10–11% oxygen (w/w), thereby enhancing the combustion process in an engine. It has also been reported that the use of tertiary fatty amines and amides can be effective in enhancing the ignition quality of the biodiesel without having any negative effect on its cold flow properties. However, starting problems persist in cold conditions. Further, biodiesel has low volumetric heating values (about 12%), a high cetane number and a high flash point. The cloud points and flash points of biodiesel are 15–25 °C higher than those of diesel.

II. EXPERIMENTAL SECTION

2.1 Process of biodiesel production

Transesterification of vegetable oils with simple alcohol has long been the preferred method for producing biodiesel. In general, there are two methods of transesterification. One method simply uses a catalyst and the other is without a catalyst. The former method has a long history of development and the biodiesel produced by this method is now available in North America, Japan and some western European countries.

2.2 Chemicals Required :

Sodium Hydroxide Pellets, Methanol Analar Grade, Waste Vegetable oil.

2.3 Apparatus Required :

Beaker (500ml), Magnetic stirrer with heating arrangement, Thermometer, Separating Funnel, Funnel, Filter paper, Holding stand for Separating Funnel, Holding stand for Funnel, 250 ml Standard Flask, Glass Rod

III. RESULTS AND DISCUSSION

3.1 Comparison Of Properties Of Biodiesel With Waste Vegetable Oil And Diesel

The properties of Biodiesel obtained from waste vegetable oil are compared with that of the oil itself and also with the commercial diesel as given in Tables 3.1

Table 3.1: Properties

Properties	waste vegetable oil	Bio diesel from waste vegetable	Commercial Diesel Fuel:
Moisture by D & S method	0.2%	0.1%	-

2.4 Procedure :

Waste vegetable oil was Collected, centrifuged and filtered to remove burned food bits, etc. Preheating was done to remove unwanted moisture present in the oil. The used oil is heated to 120 °C to remove all water present in the oil. The presence of water will decrease the yield of biodiesel and result in the formation of large amount of soap.

2.5 Preparation of Sodium Methoxide

Sodium methoxide was prepared by dissolving 4.5 gms of sodium hydroxide pellets in 250 ml of methanol in a 250 ml standard flask and shaking well. Since it is an exothermic reaction heat will be produced and it should be allowed to cool before mixing with vegetable oil.



2.6 Biodiesel Production

The cleaned vegetable oil is taken in a 500ml beaker and placed on a magnetic stirrer cum heater. The oil was stirred and heated to a required temperature. After attaining required temperature, the methoxide prepared was added gradually while stirring. The temperature was maintained at constant value for 2 hrs and it is transferred into a separating funnel and allowed to settle for over night. After settling we can find two layers of product, upper layer is biodiesel and the bottom layer is glycerin. The bio diesel thus formed was separated and washed to remove soap present in the biodiesel. Washing is continued till you get a clear biodiesel. The biodiesel was analysed for its physical properties as well as for the fuel characteristics for the engine.

Once the glycerin and biodiesel phase have been separated, the excess reactant (methanol) and the catalyst (NaOH) is removed by simply heating and also washing continuously with water.

2.7 Methyl Ester Washing

Once separated by glycerin the biodiesel is purified by gently washing with warm water to remove residual catalyst or soaps, dried and sent to storage. This is normally the end of the production process resulting in a clear amber-yellow liquid with a viscosity to petroleum. Since the ester has an extremely high tendency to form an emulsion on contact with water, care must be taken to maintain low levels of mixing during washing.

Inorganic acidity	Nil	Nil	-
Acidity as mg KOH/gm oil	0.32	0.12	-
Density at 15 °C (gm/cc)	0.9240	0.8966	0.84
Ash content	0.006%	0.004%	-
Kinematic viscosity at 40°C (cSt)	27.2	6.3	4 – 5
Flash Point by TMCC	212 °C	196 °C	76 °C
Fire Point by TMCC	220 °C	204 °C	80 °C
Pour point by TMCC	Below -11 °C	-11 °C	-16 °C
Insoluble in hexane	0.008%	0.006%	-
Gross calorific value (kcal/kg)	10,125	10,425	10,900
Sulfur content	0.09%	0.06%	45
Carbon residue	0.46%	0.33%	-
Saponification Value	188	-	-
Iodine Value	141.5		

3.2 Effect Of Operating Parameters On Biodiesel Yield

3.2.1 Effect of Temperature

The low level of temperature was chosen as the room temperature and the high level was chosen as 105 °C. Higher temperature not only decreased the time required to reach maximum conversion but also the cost of energy for heating the apparatus would not exceed the value of time saved. Therefore, high temperature of 85 °C is considered to be the optimum temperature for conversion. The pressure is maintained at atmospheric level using the condenser.

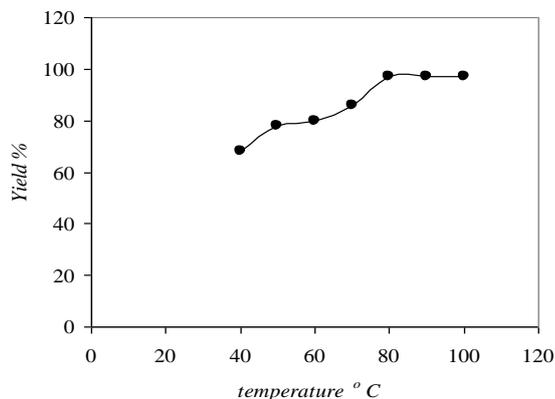


Fig 3.1 Effect of temperature on Biodiesel Yield

3.2.2 Effect of concentration of Methanol

The next step of the experimental design was used to determine the optimum values of methanol concentration. Inspection of result shows that concentration of methanol is the most important independent factor that affects the degree of conversion. Increasing the methanol concentration up to 100 % excess than the stoichiometric proportion will yield an optimum conversion. There is a higher yield at higher concentration of methanol but the energy required for the recovery of methanol becomes higher.

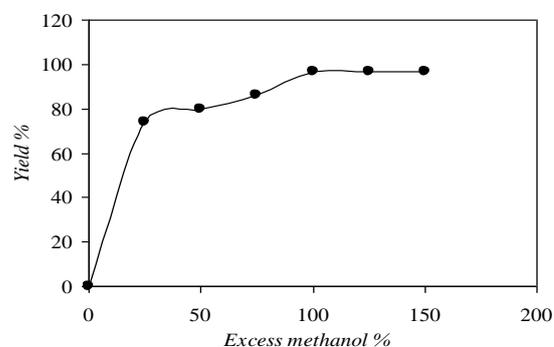


Fig 3.2 Effect of excess methanol on Biodiesel Yield

3.2.3 Effect of Catalyst Concentration

The next in optimizing the process for production of Biodiesel was to study the effect of catalyst concentration on the biodiesel yield. It is clearly shown that the biodiesel yield increases and then reaches the optimum conversion at 5-6% of the weight of the catalyst. Hence, this value was chosen for the production of biodiesel from the waste oils.

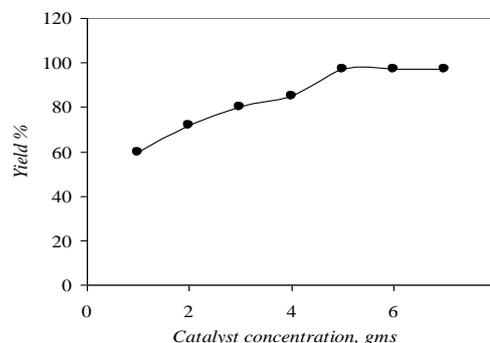


Fig 3.3 Effect of catalyst concentration on Biodiesel Yield

IV. CONCLUSION

Waste vegetable oil was transesterified using methanol in presence of alkali (NaOH) and the biodiesel obtained was studied for fuel properties. The effect of temperature, catalyst concentration and the concentration of methanol on yield of biodiesel were studied and the

optimum values of the above variables are fixed for the high yield. The properties of bio diesel was compared with the diesel. The calorific value, flash & fire point slightly varies with that of diesel. The performance characteristics of diesel engine is being carried out.

REFERENCES

- [1]. Ministry of Non-conventional Energy Sources (MNES) and Planning Commission; 2002.
- [2]. Goering CE, Schwab AW, Daugherty MJ, PrydeEH, Heakin AJ. 'Fuel properties of eleven vegetable oils' Trans ASAE 1982;85:1472– 83.
- [3]. Marckley KS. 'Fatty acids', 2nd ed. New York Interscience; 1960.
- [4]. SEA News Circular. The Solvent Extractors Association of India, vol. X; 1996.
- [5].
- [6]. Otera J. 'Transesterification'. Chem Rev 1993;93(4):1449–70.
- [7]. Freedman B, Butterfield RO, Pryde EH. 'Transesterification kinetics of soybean oil' J Am. Oil Chem Soc 1986;63(10):1375–80.
- [8]. Nouredini H, Zhu D. 'Kinetics of transesterification of soybean oil' J Am Oil Chem Soc 1997;74(11), 1457–63.
- [9]. Nye MJ, Southwell PH. 'Esters from rapeseed oil as diesel fuel' Proceedings of Vegetable Oil as Diesel Fuel—Seminar III, Peoria, IL; 1983.
- [10]. Agarwal AK. 'Vegetable oils versus diesel fuel: development and use of biodiesel in a compression ignition engine. TERI Inf Dig Energy (TIDE) 1988;8:191–203.
- [11]. Freedman B, Pryde EH, Mounts TL. 'Variables affecting the yields of fatty esters from transesterified vegetable oils' J Am Oil Chem Soc 1984;61(10):1638–43.
- [12]. Feuge RO, Gros AT. 'Modification of vegetable oils. VII Alkali catalyzed interesterification of pea- nut oil with ethanol'. J Am Oil Chem Soc 1949;26(3):97.
- [13]. Hui YH, editor. 'Bailey's industrial oil fats: industrial and consumer non edible products from oils and fats' New York: Wiley; 1996. p. 5.
- [14]. Fillieres R, Benjelloun-Mlayah B, Delmas M. 'Ethanolysis of rapeseed oil: quantification of ethylesters, mono-, di-,and triglycerides Krawczyk.T. 'Bio glycerol by high performance size-exclusion chromatography'. J Am Oil Chem Soc 1995;72(4):427–32.
- [15]. Formo MW. 'Ester reactions of fatty materials'. J Am Oil Chem Soc 1954;31(11):548–59.
- [16]. Saka S, Dadan K. 'Biodiesel fuel, from rapeseed oil as prepared in supercritical methanol'. Fuel 2001;80:225.

A Two Input Single Output Z-Sourced Dc-Dc Converter for Renewable Applications

Valluri Satya Srinivas¹, E.Vargil Kumar², K.Bhavya³
^{1, 2, 3}, (EEE, Gudlavaleru college of Engg/ Jntuk, India)

Abstract: This paper proposes a double input dc-dc converter based on Z-source converters. In the proposed converter, the input dc voltage can be boosted and also input dc sources can deliver power to the load individually or simultaneously, so combination of a battery with one of the new energy sources such as solar array, wind turbine or fuel cell can be used as input sources. A new control technique individual channel designing is proposed in this to control input supplies of the circuit finally, the simulation results are presented to confirm the theoretical analysis.

Keywords: Z-source converter; double input; dc-dc converter.

I. INTRODUCTION

The renewable energy such as photovoltaic (PV) and wind has created various electric energy sources with different electrical characteristics for the modern power system. In order to combine more than one energy source, such as the solar array, wind turbine, fuel cell (FC) and commercial ac line to get the regulated output voltage, the different topologies of multi input converters (MICs) have been proposed in recent years [1]-[4]. Traditionally, two dc voltage sources are connected to two independent dc-dc power converters to obtain two stable and equivalent output voltages, which are then connected to the dc bus, to provide the electric energy demanded by the load. Another approach for the double-input dc-dc converter is to put two dc sources in series to form a single voltage source where traditional dc-dc power converters can be used to transfer power to the load. In order to transfer power individually, each dc voltage source needs a controllable switch to provide a bypass short circuit for the input current of the other dc voltage source to deliver electric energy continuously [3], [4]. Another approach is to put PWM converters in parallel with or without electrical isolation using the coupled transformer [5]. Control schemes for these MICs with paralleled dc sources are based on the time-sharing concept because of the clamped voltage. Because of the voltage amplitude differences between two dc sources, only one of them can be connected to the input terminal of the dc-dc converter and transfer power to the load at a time [3]. The general form of a MIC consists of several input sources and a single load, conceptually shown in Fig. 1. In general, all of the input sources can deliver power to the load. When only one of the input sources feeds the MIC, it will transfer power to the load individually and the MIC will separate as a PWM Converter.

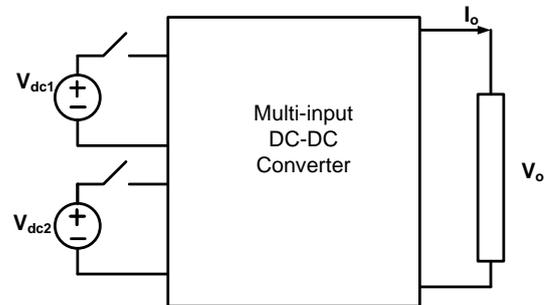


Figure 1. General form of multi input converter

In other words, when more than one input sources are supplied to the MIC, all input sources will deliver power to the load simultaneously without disturbing each other's operation. The objective of this paper is to propose a double-input dc-dc converter which has the following advantages: The dc sources can deliver power to the load individually or simultaneously; the multi winding transformer is not needed; the magnitude of the input dc voltage can be higher or lower than the one with a regulated output; minimum switching devices are used in the converter circuit. The proposed double input dc-dc converter is proper for renewable-energy applications and combination of two different sources (such as battery and photovoltaic or fuel cell).

The purpose of this paper is to introduce the novel Individual Channel Design MATLAB Toolbox and to provide an explanation of how it should be used. At present the software considers only 2x2 control systems. The toolbox is a valuable aid for analysing and designing multivariable control systems. Given a set of specifications for a 2x2 multivariable control system the appropriate use of the toolbox can lead to successful controllers. The process is based on an iterative procedure. Closed loop simulations (in SIMULINK®) are included so results can be tested. Final stability margins and robustness measures are also assessed. The toolbox is inspired on a new approach for multivariable control systems, referred as Individual Channel Design (ICD). ICD is a novel analytical framework that allows the analysis and synthesis of Multi-variable control systems under the context of the Multivariable Structure Function (MSF) by applying classical techniques based on the Bode and Nyquist plots. With the help of this framework it is possible to investigate the potential and limitations for feedback design of any multivariable linear time-invariant control system. Although ICD is in principle a feedback structure based on diagonal controllers, it can be applied to

any cross coupled multivariable system . It is based on the definition of individual transmission channels. In this context the control design is an interactive process that involves the required specifications, plant characteristics, and the multivariable feedback design process itself. Once the channels are defined it is possible to form a feedback loop with the compensator specially designed to meet customer specifications. In this manner the multivariable control design problem is reduced to the design of a single–input single–output control for each channel. ICD has been reported in some control strategies, such asin small scale power networks with embedded generation, in the automotive and the aerospace industry. So far, this toolbox has been used in different control tasks, from induction motors , synchronous generators , to submarines . It has also been employed to provide a design example in a family of plants that range from simple and decoupled to a highly coupled and non– minimum phase one. The document is organised as follows. Section II includes a brief review of ICD. In fact, it is highly recommended to have some expertise in the subject in order to exploit the toolbox successfully. In Section III the design procedure is explained. An introduction to the software and a design example using the toolbox is carried out in Section IV. It is shown in full detail in order to provide a clear demonstration of the toolbox potential. The example here presented corresponds to an interesting and rather challenging design case . Finally, in Section V the conclusions end the paper.

II. CIRCUIT CONFIGURATION AND OPERATION PRINCIPLE OF THE PROPOSED CONVERTER

A.Z-Source Converters

Z-source converters are modern group of power electronic converters which can overcome problems with traditional converters. The Z-source inverter is a novel topology [6] that overcomes the conceptual and theoretical barriers and limitations of the traditional voltage-source converter and current-source converter. The concept of Z-source was used in direct ac-ac power conversion [7]. Similarly, the concept of Zsource also was extended to dc-dc power conversion [8].

B. Circuit Configuration of Proposed Converter

The schematic circuit diagram of the proposed double-input Z-source dc-dc converter with two different voltage sources is shown in Fig. 2. It consists of two different input sources, Vdc1 and Vdc2, and four diodes, S1-S4, applied to provide current path in different states. In this paper, permanent connection of input dc sources is considered, so S1 and S2 can be replaced with active switches if it's required to connect and disconnect each of

sources to input side of converter frequently. Energy receiver, converter and transmitter sections are situated in the middle side of the converter. This section is a two-port network that consists of a split-inductor L1 and L2 and capacitors C1 and C2 connected in x-shape which is named "Z-network". An active switch, S, is situated in output port of Z-network to control input and output power of converter. The final section of converter is a LC filter beside the load in order to reject output signal ripple.

C. Principle Operation of Double Input DC-DC Converter

There are four different operation states with respect to active or inactive states of dc sources. As previously mentioned, both of the input sources can deliver power to the load either individually or simultaneously through the MIC. When only one of the input sources feeds the MIC, it transfers power to the load individually and the MIC will operate as does a PWM converter. Table I summarizes the operation states of the proposed double input dc-dc converter.

D. State 1, both source 1 and source 2 are active

Fig. 3 shows equivalent circuit of this state. When both source1 and source 2 are active, the converter input dc voltage is sum of voltage of two series dc sources, as Fig. 3 and (1) illustrate.

$$V_{in} = V_{dc1} + V_{dc2} \quad (1)$$

In this state, because both two sources are active, S1 and S2 are forward biased and S3 and D4 are reverse biased. Thus, the sources current enters in Z-network through S1 and S2 and after passing load impedance, comes back into sources through negative polarity.

E. State 2, source 1 is active and source 2 is inactive

The equivalent circuit of this state is shown in Fig. 4. In this state, source 1 is active, so only this source provides converter (consequently load) energy. Because of source 1 is active then S1 is forward biased and S3 is reverse biased, so current follows from S1 to Z-network to load.

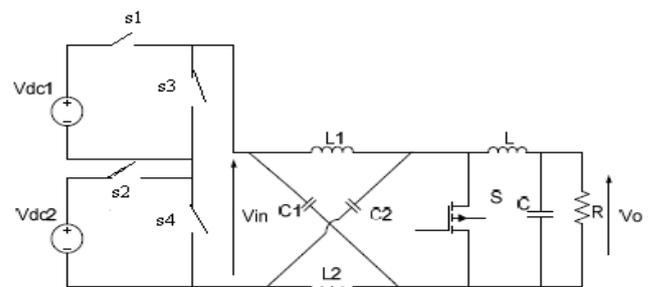


Fig.1 double input z source converter

TABLE I. STATES OF DOUBLE INPUT DC-DC CONVERTER

state	Sources States		Switches States				V_{in}
	V_{dc1}	V_{dc2}	s1	s2	s3	s4	
1	Active	Active	On	On	Off	Off	$V_{dc1} + V_{dc2}$
2	Active	inactive	On	Off	Off	On	V_{dc1}
3	inactive	Active	Off	On	On	Off	V_{dc2}
4	inactive	inactive	Off	Off	On	On	0

In reverse path from load to the source, current can't pass through source 2 and S2, so S4 is forcedly turned on and conduct current to source 1. In this state, converter input dc voltage is only provided by source 1, as (2) shows.

$$V_{in} = V_{dc1} \quad (2)$$

F. State 3, source 1 is inactive and source 2 is active

If source 1 is eliminated for each reason and source 2 is active, the converter can operate normally without effect of source 1 elimination. Fig. 5 shows the equivalent circuit for This state. In state 3, it's only source 2 that supplies converter

And load. Source 2 activation causes forward bias of S2 and Reverse bias of S4. Because of source 1 disconnection, current Passes through S3 and indeed, current turns it on forcedly to Complete current path. In this state, converter input dc voltage is only provided by source 2, as (3) shows.

$$V_{in} = V_{dc2} \quad (3)$$

G. State 4, both source 1 and source 2 are inactive

Basically, this state is only following of one of the previously mentioned three states. Because in this state both dc sources are inactive and disconnected from converter, D1 and D2 are forcedly turned off and consequently, the only existing path for remain current, from previous state, is provided by D3 and D4. Thereupon, in state4 D3 and D4 are turned on. Fig. 6 shows equivalent circuit of this state. Input voltage is zero in this state as shown in (4).

$$V_{in} = 0 \quad (4)$$

Obviously, because both dc sources disconnect from converter, duration of this state is very short and when current descends to zero, whole of converter will be inactive.

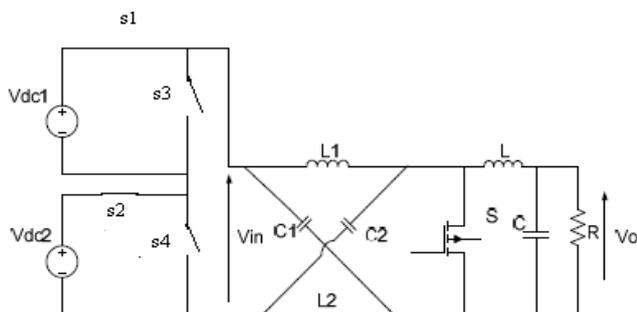


Fig.3 state 1 equivalent circuit

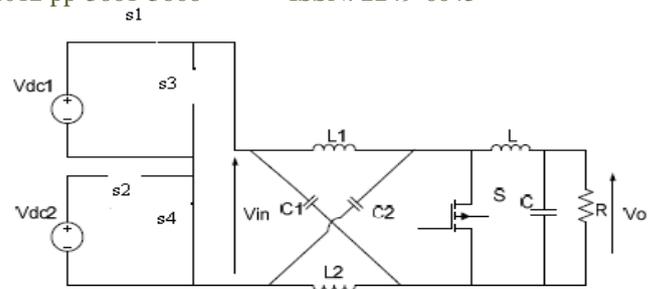


Fig.4 State 2 Equivalent circuit

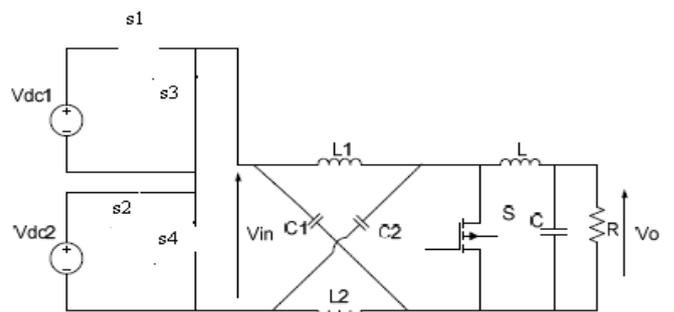


Fig.5 State 3 Equivalent circuit

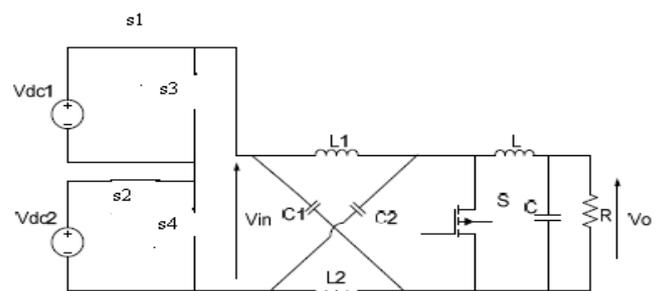


Fig6. State 4 Equivalent circuit

III. STEADY STATE ANALYSIS OF DOUBLE INPUT Z-SOURCE DC-DC CONVERTER

Assuming that the inductors L1 and L2 and capacitors C1 and C2 have the same inductance(L) and capacitance(C), respectively, the Z-source network becomes symmetrical. From the symmetry and the equivalent circuits, we have

$$V_{C1} = V_{C2} = V_C ; V_{L1} = V_{L2} = V_L \quad (1)$$

Given that the inverter bridge is in the shoot-through zero state for an interval of , during a switching cycle T_0 , and from the equivalent circuit, Fig. 6, one has

$$V_L = V_C \quad V_d = 2V_C \quad V_i = 0 \quad (2)$$

Now consider that the inverter bridge is in one of the eight nonshoot-through states for an interval of T_1 , during the switching cycle, . From the equivalent circuit, Fig. 7, one has

$$\begin{aligned} V_L &= V_o - V_C ; v_d = V_o \\ v_i &= V_C - v_L = 2V_C - V_o \end{aligned} \quad (3)$$

Where V_0 is the dc source voltage and $T=T_0+T_1$. The average voltage of the inductors over one switching period(1) should be zero in steady state, from (2) and (3), thus, we have

$$V_L = \bar{v}_i = \frac{T_0 \cdot V_C + T_1 \cdot (V_o - V_C)}{T} = 0 \quad (4)$$

$$\frac{V_C}{V_o} = \frac{T_1}{T_1 - T_0} \quad (5)$$

Similarly, the average dc-link voltage across the inverter bridge can be found as follows:

$$V_i = \bar{v}_i = \frac{T_0 \cdot 0 + T_1 \cdot (2V_C - V_o)}{T} = \frac{T_1}{T_1 - T_0} V_o = V_C \quad (6)$$

The peak dc-link voltage across the inverter bridge is expressed in (3) and can be rewritten as

$$\hat{v}_i = V_C - v_L = 2V_C - V_o = \frac{T}{T_1 - T_0} V_o = B V_o \quad (7)$$

$$B = \frac{T}{T_1 - T_0} = \frac{1}{1 - 2\frac{T_0}{T}} \geq 1 \quad (8)$$

is the boost factor resulting from the shoot-through zero state.

The peak dc-link voltage v_i is the equivalent dc-link voltage of the inverter. On the other side, the output peak phase voltage from the inverter can be expressed as

$$\hat{v}_{ac} = M \cdot \frac{v_i}{2} \quad (9)$$

where is the modulation index. M Using (7), (9) can be further expressed as

$$\hat{v}_{ac} = M \cdot B \cdot \frac{V_o}{2} \quad (10)$$

For the traditional V-source PWM inverter, we have the well known relationship: $v_{ac}=M \cdot V/2$. Equation (10) shows that the output voltage can be stepped up and down by choosing an appropriate buck–boost factor ,

$$B_B = M \cdot B = (0 \sim \infty) \quad (11)$$

From (1), (5) and (8), the capacitor voltage can expressed as

$$V_{C1} = V_{C2} = V_C = \frac{1 - \frac{T_0}{T}}{1 - \frac{2T_0}{T}} V_o \quad (12)$$

The buck–boost factor B_B is determined by the modulation Index M and boost factor B. The boost factor B as expressed in (8) can be controlled by duty cycle (i.e., interval ratio) of the shoot-through zero state over the non shoot-through states of the inverter PWM. Note that the shoot-through zero state does not affect the PWM control of the inverter,

because it equivalently produce the same zero voltage to the load terminal. The available shoot through period is limited by the zero-state period that is determined by the modulation index.

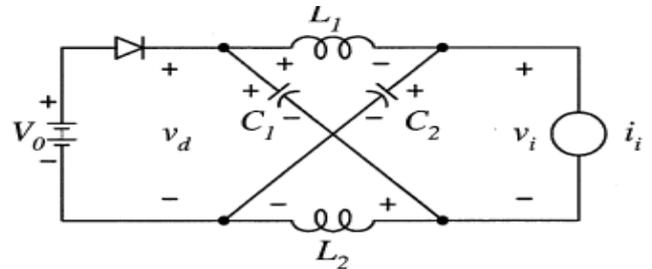


Fig. 7. Equivalent circuit of the Z-source inverter viewed from the dc link.

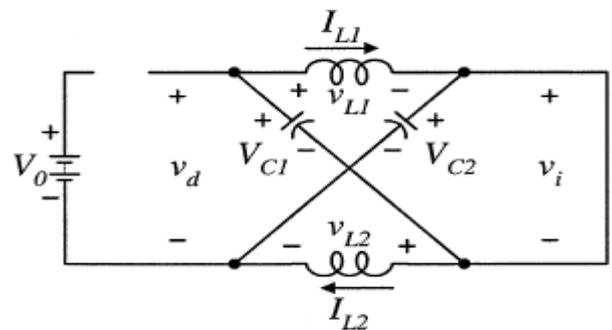


Fig. 8. Equivalent circuit of the Z-source converter viewed from the dc link when the inverter bridge is in the shoot-through zero state.

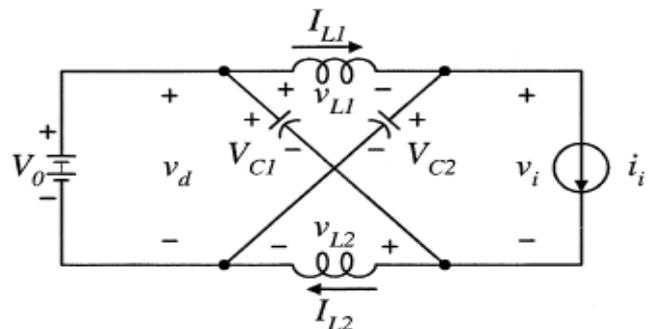


Fig. 9. Equivalent circuit of the Z-source converter viewed from the dc link when the inverter bridge is in one of the eight nonshoot-through switching states.

IV. INDIVIDUAL CHANNEL DESIGNING

In a typical control design task the performance is specified in terms of an output response to a given input. Meanwhile, in multivariable control, there are a number of inputs and outputs and, although it may be clear which inputs are intended to drive which outputs, the design task can be obscured by cross-coupling via the plant dynamics. Nevertheless, for clarity of both performance specification and design, it remains desirable to consider the inputs and outputs in pairs. The situation is depicted in Fig. 1, where \mathbf{G} is the plant and \mathbf{K} is the controller. Input ri is paired with

output y_i in accordance with specifications. An individual pairing is called a *channel*. Then, channel C_i is the pairing between r_i and y_i .

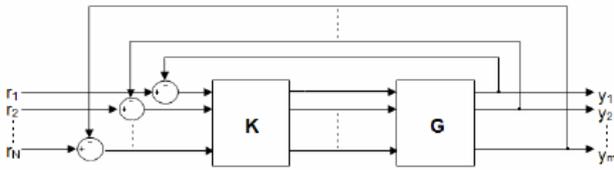


Fig.10 MIMO multivariable system. Channel definition

When the plant cross-coupling is weak, the design task reduces to a set of SISO design tasks and a scalar controller can be designed separately for each channel. In such context, the most appropriate methodology is to apply classical Nyquist/Bode analysis and design to each channel ICD is a framework in which Bode/Nyquist techniques can be applied directly to the channels not only when cross-coupling is weak but in *all circumstances* including when cross-coupling is strong. The multivariable system is decomposed into an equivalent set of SISO systems. Each SISO system is the open-loop channel transmittance between input r_i and output y_i , with the feedback loop between output y_i and input r_i open but all other feedback loops closed, for a particular choice of i . What is particular to Individual Channel Design is that the SISO channel transmittances are reformulated to make explicit the role of the plant structure. Scalar multivariable structure functions (MSFs) to which the individual channel transmittances are simply related encapsulate the significant aspects of the plant structure. The multivariable nature of the original plant is maintained in the equivalent SISO systems through the multivariable structure functions with *no loss of information*. The ICD set up for a 2-input 2-output plant is shown next for completeness. Let a 2x2 plant

$$Y(s) = G(s)U(s) \tag{1}$$

be represented by

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{2}$$

where $g_{ij}(s)$ represents scalar transfer functions, $y_i(s)$ the outputs, and $u_i(s)$ the inputs of the system, with $i, j = 1, 2$. If a diagonal controller is given by

$$u(s) = K(s) \cdot e(s) \tag{3}$$

$$\begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} = \begin{bmatrix} K_{11}(s) & K_{12}(s) \\ K_{21}(s) & K_{22}(s) \end{bmatrix} \begin{bmatrix} e_1(s) \\ e_2(s) \end{bmatrix} \tag{4}$$

with $e_i(s) = r_i(s) - y_i(s)$, where $r_i(s)$ represents the plant references, then the open loop input-output channels are clearly defined from Figs. 2 and 3 as

$$C_i(s) = k_{ii}(s)g_{ii}(s)(1 - \gamma_a h_j(s)) \tag{5}$$

where i not equal j and $i, j = 1, 2$. The complex valued function

$$\gamma_a(s) = \frac{g_{12}(s)g_{21}(s)}{g_{11}(s)g_{22}(s)} \tag{6}$$

is referred to as the multivariable structure function (MSF). The functions $h_i(s)$ are:

$$h_i(s) = \frac{k_{ii}(s)g_{ii}(s)}{1 + k_{ii}(s)g_{ii}(s)} \tag{7}$$

The interaction or cross coupling between the channels can be evaluated through a transfer function. For instance, the influence of channel- j on channel- i is

$$d_i(s) = \frac{g_{ij}(s)}{g_{jj}(s)} h_j(s) r_j(s) \tag{8}$$

It is clear that the correct interpretation of the MSF (6) is of great importance because

- It determines the dynamical characteristics of each input-output configuration;
- It has an interpretation in the frequency domain;
- Its magnitude quantifies the coupling between the channels (in the frequency domain);
- It is related to the plant transmission zeros (zeros of $1 - a(s)$, $|G(s)| = g_{11}(s)g_{22}(s) - g_{12}(s)g_{21}(s) = 0$); $a(s) = 1$ determines the non-minimum phase condition;
- Its closeness to $(1, 0)$ in the Nyquist plot indicates to
- what extent the plant is sensitive to uncertainty in terms of gain and phase margins. This fact plays a key role in order to obtain robust controllers.

A block diagram of the feedback system with the diagonal controller is shown in Fig. 2 and the equivalent scalar channels are shown in Fig. 3.

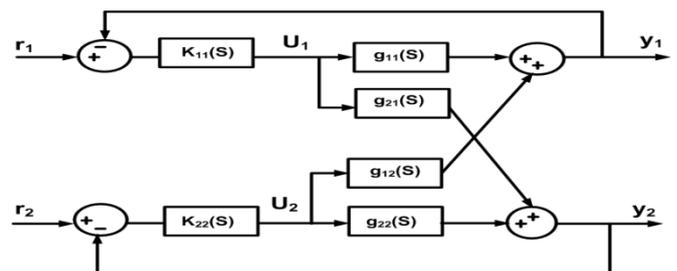


Fig.11 The 2-input 2-output multivariable system with a diagonal Controller

It should be emphasised that in the individual channel representation of the multivariable system there is no loss of information [3]. The multivariable character and cross coupling of the plant are contained in the MSF and the cross coupling terms. That is, (5)–(8) are equivalent to the closed loop matrix function

$$G_d(s) = (I + G(s)K(s))^{-1}G(s)K(s) \tag{9}$$

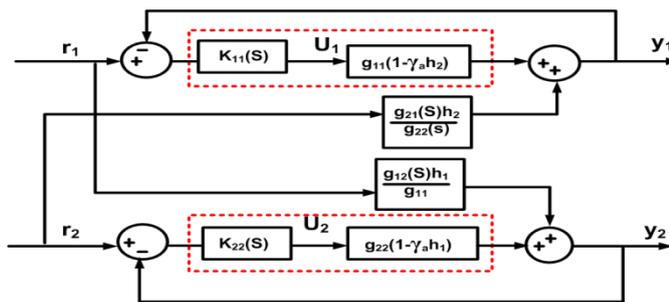


Fig.12 Equivalent channels of a 2-input 2-output control system

It can be proven that in order to stabilise (9) it is just necessary to stabilise the channels given by (5) [7,9]. In general stabilisation of the diagonal elements of $G(s)$ is not required [9]. The open loop system dynamical structure with a diagonal controller is summarised in Table I [3]. Notice that the coupling can be expressed in decibels directly from the channels (5) by means of functions $a(s)hj(s)$. On the other hand, it is possible to determine the dynamical structure of the system using Table I and analysing the Nyquist plot of $(I - a(s)hj(s))$.

Table I. Dynamical structure of open loop channels

Channel	Zeros	Poles
$C_1(s)$	Zeros of $(1-\gamma(s)h_2(s))$	Poles of $g_{11}(s), g_{12}(s), g_{21}(s), h_2(s)$
$C_2(s)$	Zeros of $(1-\gamma(s)h_1(s))$	Poles of $g_{22}(s), g_{12}(s), g_{21}(s), h_1(s)$

It is clear that the controller performance characteristics are determined by the MSF. If the transfer matrix $G(s)$ possess a non-minimum phase transmission zero, some problems will arise while stabilising it especially if the value of the zero is smaller than the desired cut-off frequency. Moreover, the robustness of the channels can be established in terms of gain and phase margins as the Nyquist paths of the functions $ya(s)hi(s)$ do not pass near $(1,0)$. Thus the design of $kii(s)$, which should provide adequate gain and phase margins for $kii(s)gii(s)$, can be obtained through an iterative process. It should be noticed that the RHPPs of the channels are RHPPs of individual transfer functions as established in Table I. On the other hand, the RHPZs of the channels are RHPZs of $(I - y a(s)hi(s))$. Moreover, the number of RHPZs of the previous function can be determined after applying the Nyquist Stability Criterion. In fact, the RHPZs of $(I - \gamma a(s)hi(s))$ are given by

$$Z = N + P \quad (10)$$

where P is the number of RHPPs of $\gamma a(s)hi(s)$ and N is the number of encirclements in clockwise direction to $(1,0)$ of the complex plane in the Nyquist diagram of $y a(s)hi(s)$.

The dynamical structure of the 2x2 plant is determined by the input-output channels defined by pairing each input to each output. For instance:

- (a) $C_1(s) : u_1(s) - y_1(s)$ with $\gamma_a(s) = g_{12}(s)g_{21}(s)g_{11}(s)g_{22}(s)$
 $C_2(s) : u_2(s) - y_2(s)$
- (b) $C_1(s) : u_1(s) - y_2(s)$ with $\gamma_b(s) = g_{11}(s)g_{22}(s)g_{12}(s)g_{21}(s)$
 $C_2(s) : u_2(s) - y_1(s)$

The coupling characteristic of each configuration is determined from $\gamma a(s)$ and $\gamma b(s)$ –their associated MSFs.

V. SIMULATION RESULTS

Simulation of double input Z-source dc-dc converter was performed using MATLAB/SIMULINK to confirm above analysis. Simulation consists of double input circuit with ICD control technique by using z source inverter in this the z source inverter output depends upon the ICD design the because of this reason there is only one operation will be performed which best for the output voltage Converter parameters in the simulation were as in Table II.

TABLE II. Simulation Parameters

Parameter	Value
V_{dc1}	40 V
V_{dc2}	100 V
R	15 Ω
$C1=C2$	1000 μ F
C	500 μ F
$L1=L2=L$	0.5Mh
Switching Frequency	10kHz
Duty ratio (D)	30%

Here the input of double input z source converter are taken as wind and solar energies .for solar input maximum power is achieved by using MPPT algorithm. For wind without MPPT supply is produced to the z source inverter .based upon the out put voltage of load the two inputs are produced for z source converter.

The input of the solar panels as shoen infig.13.In that circuit w are going to apply ICD circit which are taken transfer function as shown beow.the solar out pur currents are shown fig.13.the soalr out put will vry depending upon the required input voltage of the z source converter as well here shown the speed & torque of the wind generator in fig 14.

$$K_{11} = \frac{2}{s^2 + 2s + 1} \quad K_{22} = \frac{3}{s^2 + 2s + 1}$$

$$g_{11} = \frac{2}{s^2 + 2s + 1} \quad g_{12} = \frac{-2}{s + 1}$$

$$g_{21} = \frac{-1}{s^2 + 2s + 1} \quad g_{22} = \frac{6}{s^2 + 2s + 6}$$

approximately zero.the ICD control block is implemented to control the switches at the z at source converter ..

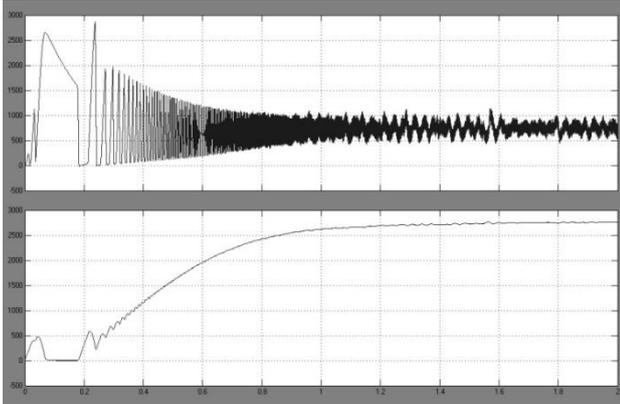
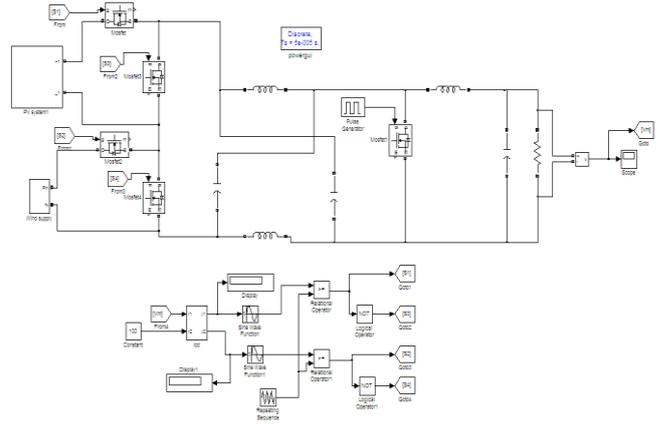


Fig.13a. solar output currents when mppt is not applied & 13b. When mppt is applied

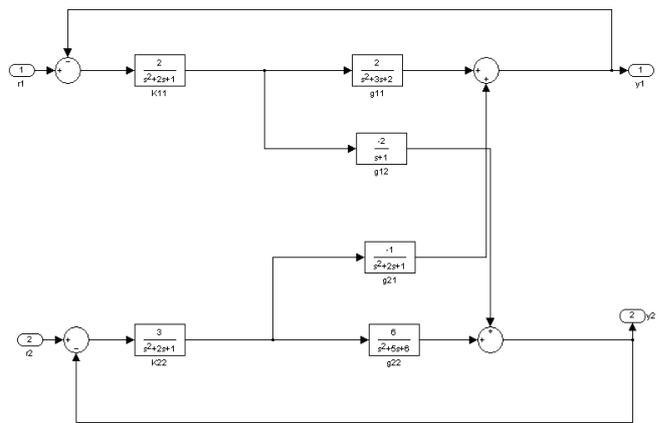


Fig.15 MATLAB simulation of proposed converter along with diagram of the ICD controller

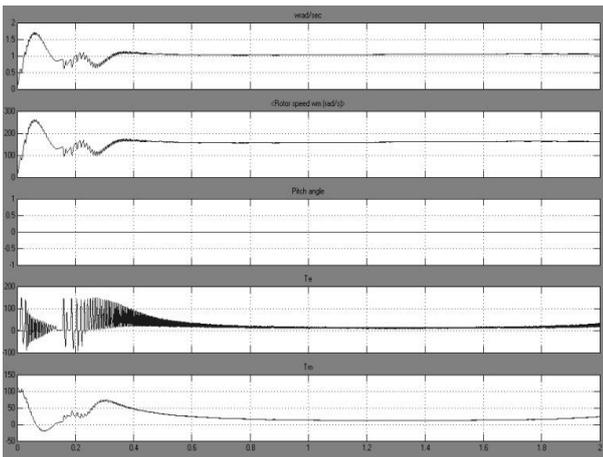


Fig.14 wind speed ,rotor speed, pitch angle, electrical torque .mechanical torque

Here icd control techniqe is very effiecient technmique.the ICD block diagram i fig.15 .Solar iout pur current is intally is 2700amps but according to requirement we need 800 amps.after applying the MPPT the output of the solar is reduced to nearly 800.wind supply is basecupon the wind speed .rotor speed is given as 150rad/sec and also wind speed is 1rad/sec.the pitch angle is maintained zero always. Electromagnetic torque & mechanical torque maintained

The bode plots of the controller is implemented by the MIMO tool box in MATLAB. The bode plots of the controller.

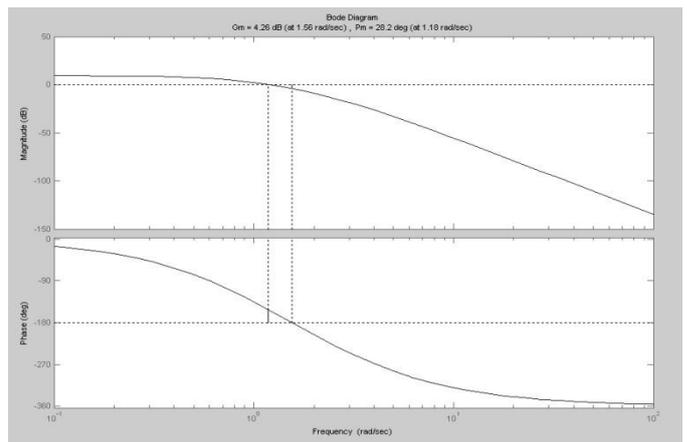


Fig.16(a) bode plot of the K11& K2

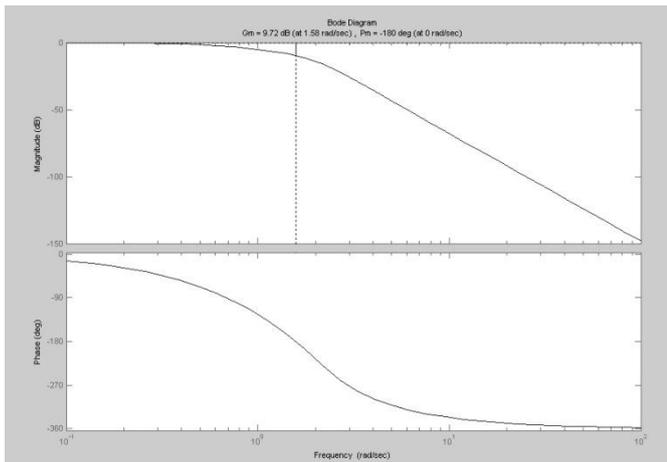


Fig.16(b) bode plot of the K11& K2

Bode gamma plots of the controllers also shown in the fig.17. the performance of the controller is shown in the bode plots to draw the plots we implanted a ICD code which is call as ICD tool is used in MIMO tools.

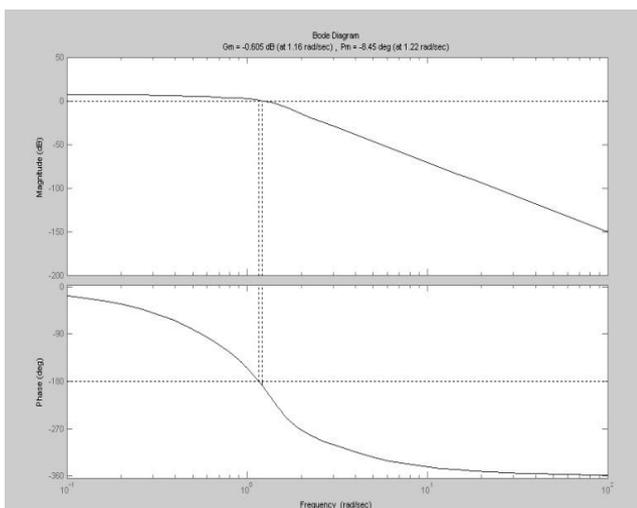


Fig.

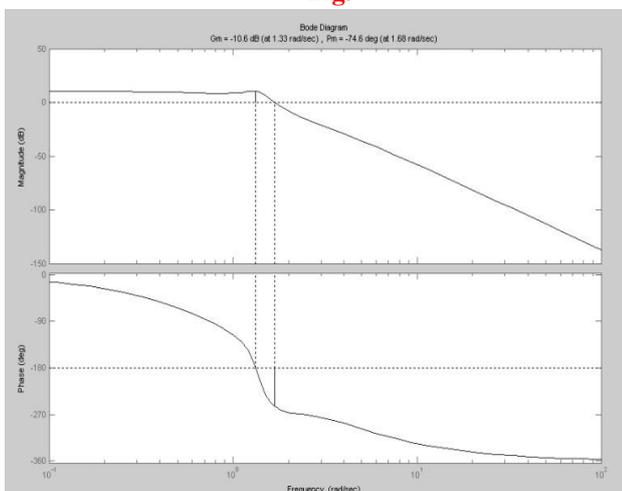


Fig.17 Bode plot gamma of K11 & K22

Finally the output of the is shown in fig.18.the output of te z sources converter is 50V hence the out put will be almost constant after certain time.

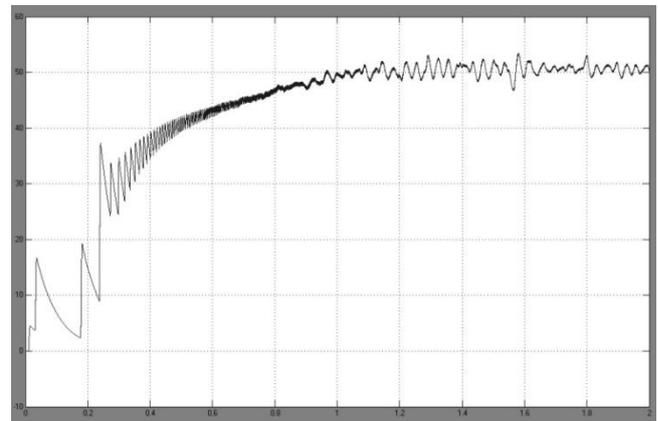


Fig.18 output voltage of the Z-source converter

VI. CONCLUSION

In this paper, double input Z-source dc-dc converter is proposed. The operation principle, ICD controller implementation and its operation and steady-state analysis is explained in detail. The analysis and simulation results show the input dc sources can deliver power to the load, as failure of each input sources doesn't disturb the other's operation.

Two input sources can have different characteristics and voltage. Also, converter controls output power with only one active switch which can reduce cost and improve the reliability. Boosting feature of converter makes it proper for new energy applications.

REFERENCES

- [1] E. Muljadi and H. E. McKenna, "Power quality issues in a hybrid power system," *IEEE Trans. Ind. Appl.*, vol. 38, no. 3, pp. 803-809, May/June 2002.
- [2] F. Giraud and Z. M. Salameh, "Steady-state performance of a gridconnected rooftop hybrid wind-photovoltaic power system with battery storage," *IEEE Trans. Energy Convers.*, vol. 16, no. 1, pp. 1-7, Mar 2001.
- [3] Yaow-Ming Chen, Yuan-Chuan Liu, and Sheng-sien Lin, "Double- Input PWM DC-DC Converter for High-/Low-Voltage Sources," *IEEE Trans. Ind. Electron.*, vol. 53, no. 5, October 2006.
- [4] Yuan-Chuan Liu and Yaow-Ming Chen, "A systematic approach to synthesizing multi-input dc-dc converters" *IEEE Trans. Power Electron.*, vol. 24, no. 1, January 2009.
- [5] H. Matsuo, W. Lin, F. Kurokawa, T. Shigemizu, and N. Watanabe, "Characteristics of the multiple-input dc-dc converter," *IEEE Trans. Ind. Electron.*, vol. 51, no. 3, pp. 625-631, Jun 2004.

- [6] F. Z. Peng, "Z-Source inverter," *IEEE Trans. Ind. Appl.*, vol. 39, no. 2, pp. 504-510, March/April 2003.
- [7] Xu Peng Fang, Zhao Ming Qian, and Fang Zheng Peng, "Single-phase Z-Source PWM AC-AC Converters," *IEEE Power Electron. Lett.*, vol. 3, no. 4, December 2005.
- [8] E. Licéaga-Castro, C.E. Ugalde-Loo, J. Licéaga-Castro, and P. Ponce. "An efficient controller for SV-PWM VSI based on the multivariable structure function". To be presented at: *44th IEEE CDC-ECC*, Spain, 2005.

Investigation and Analysis of Vulnerability of Attacks on Watermarked Image and Its Enhancement

G. Srinivas Reddy¹, T.Venkat Narayana Rao², Dr.A.Govardhan³

¹, Assistant Professor, Mahatma Gandhi Institute of Technology [MGIT], Hyderabad, A.P. India

², Scholar JNTU-K and Hyderabad Institute of Technology and Management [HITAM], Hyderabad, A.P, India

³, Director of Evaluation and Professor, C.S.E, Jawaharlal Nehru Technological University, Hyderabad, A.P. India

ABSTRACT: This paper demonstrate the attempt to evolve feasible solution to the scenario in which different possible attacks on embedded watermark can be avoided along with the enhancement of the content retrieved from noisy image post watermarking. With the revolution of information technology and wide area networking, data has become less private where in the admittance of media as well as the attempts to change and manipulate the contents of information has become a universal issue. Watermarking techniques have to be used to protect the copyright of the media and for the digital management but without compromising on the visual front. A universal DWT technique is used in this work. The algorithm for embedding watermark into the original image has been developed with novelty for better results than existing mechanisms. The various attacks such as image resizing, image cropping and image filtering are employed on the watermarked image to investigate the reliability of embedding algorithm. It is reported that Peak Signal to Noise Ratio (PSNR) value obtained before and after the attack has been in the ratio of 0.01 to 0.07. The extracted watermark is addressed again to enhance the PSNR value with various de-blurring techniques such as DCT compression, image cropping with noise treatment, Normal and Wiener filtering.

Keywords: Attacks, Watermarking, Image Filtering, Extraction, PSNR.

I. INTRODUCTION

The growth of networked multimedia systems has magnified the need for image copyright protection from any illegal duplication of their data and manuscripts [1]. Some serious work needs to be done in order to maintain the availability of multimedia information but, in the meantime, the industry must come up with the ways to protect intellectual property i.e. the stake of distributors and owners of data [11]. The rapid expansion of Internet in the past years has increased the availability of digital data such as audio, images and videos to the public. The idea of robust watermarking of images is to embed information data within the image with an insensible form for human visual system but in a way that protects from attacks such as common image processing operations [2]. The goal is to produce an image that looks exactly the same to a human eye but still allows its positive identification in comparison with the owner's key if necessary [12]. The fundamental tool of DWT is used in

the present study and is shown in the figure1. The number of vanishing moments describes this property. Scaling function derived from wavelets with higher number of vanishing moments can independently represent polynomials of higher orders [7]. The basic figure 1 depicts

the DWT that attracts a universal utility by the researchers in the field of image processing.

II. PROPOSED METHODOLOGY

The Phases In This Proposal Of The Present Investigation Are Mainly Centered At Attacks Possible On Embedded Watermark Image And Developing The Extracted Water Mark Image. Different Attacks May Be Possible On The Information Embedded In The Image. The Investigation Of All Possible Attacks Has Been Addressed By Using Operations Like Image Filtering, Image Resizing And Image Cropping. The Algorithm Used For Watermark Extraction Is Idwt Algorithm And Watermark Retrieved Image Is De-Blurred With Techniques Such As Normal Filtering, Wiener Filtering And Dct Compression. The Psnr Values Have Been Computed At Various Levels For Comparison And Analysis On Output Images [8].

The Haar wavelet transform which is one of the basic tools in embedding mechanism of water mark is one kind of wavelet transform implemented in the investigation as shown in figure 2-3. A digital image I with m x n pixels is transformed to the DWT frequency domain as follows. First, a cover image is decomposed into a low frequency band LL1 and three high frequency bands LH1, HL1 and HH1. Later applying the DWT on the low frequency band LL1 again will generate four lower-resolution sub-bands LL2, LH2, HL2 and HH2. This process is continued an arbitrary number of times, which is usually determined by the application available or a simple algorithm. The approximate image band LL holds the most important information of the original image. The LH, HL and HH bands contain some high-frequency information about the edge components of the signal [4]. In addition, from these DWT coefficients; the original signal can be reconstructed. This reconstruction process is called the inverse DWT (IDWT). An image can be decomposed into a pyramid structure as shown below. Figure 4 shows the image "Lena" and the transformed result after the two-level DWT transformation.

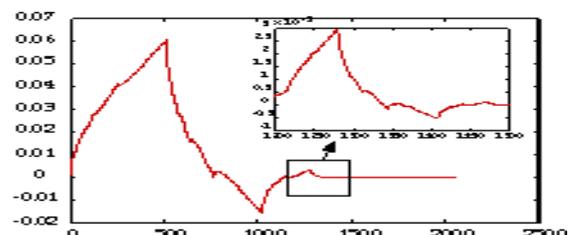


Fig. 1: Discrete Wavelet Transform

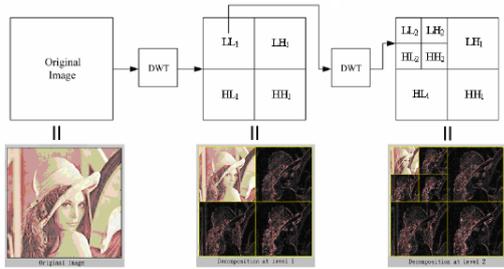


Fig. 2: Three levels Haar Decomposition

The following equations that are universally established in the literature have been employed in the investigation to generate the components of LL_1 , LH_1 , HL_1 and HH_1 values respectively.

$$LL_1(x, y) = \frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 I(2x+i, 2y+j) \quad (1)$$

$$LH_1(x, y) = \frac{1}{4} \sum_{i=0}^1 I(2x+i, 2y) - \frac{1}{4} \sum_{i=0}^1 I(2x+i, 2y+1) \quad (2)$$

$$HL_1(x, y) = \frac{1}{4} \sum_{j=0}^1 I(2x, 2y+j) - \frac{1}{4} \sum_{j=0}^1 I(2x+1, 2y+j) \quad (3)$$

$$HH_1(x, y) = \frac{1}{4} \{I(2x, 2y) + I(2x+1, 2y+1) - I(2x+1, 2y) - I(2x, 2y+1)\} \quad (4)$$

Watermarking in the DWT domain includes two parts i.e. Encoding and Decoding. In decoding method we propose hierarchical approach. Post decomposition the received image and the original image is compared. Later the signature is added in the HH_1 band and the difference of the DWT coefficients in HH_1 bands calculated with their cross correlations. In watermarking process, the image is decomposed into frequency bands using three resolutions of Haar wavelets. Figure 2 represents the idea of the octave-band structure of Haar wavelets, which offer pyramid structure [9]. We must focus on, sampling operation after every filtering. It must be understood that the choice of the Haar wavelet in our system is made for simplicity. However, we intend to investigate the influence of the selection of wavelet function in our results but, in order to test the robustness openly, we had to relinquish the idea in support of the addition of extra robustness testing procedures [14].

III. IMPLEMENTATION

3.1 Watermark Generation and Xie's DWT, Quantization, Blind Image Watermarking Algorithm

The mark is a Gaussian sequence of pseudo-random real numbers and will be denoted $X = x_1, x_2, \dots, x_n$ where n is length of watermark. The choice of the watermark length n determines to which degree the watermark is spread out. In most cases the larger the watermark the lesser the embedding strength. There is no one watermark length n that is suitable for all images, therefore it is image specific [3]. This algorithm is the stronger of the two watermarking algorithms used and was first introduced in the paper by Lihua Xie and Gonzalo R. Arce i.e. Joint wavelet compression and authentication watermarking that describes a blind watermarking algorithm for embedding watermark for authentication[13]. The watermark algorithm

is implemented in the Discrete Wavelet Transform, DWT. Xie and others also converse about implementation of SPIHT compression algorithms. Since this is a blind algorithm the watermark is extracted without the original image. The median of the sliding window is determined and quantized to obtain a reconstruction point [10]. The bit value is then determined from the associated reconstruction point which is assigned to x_i^* where X^* is extracted watermark. The extracted watermark, X^* , is then compared with that of the original watermark X . The basic mechanism for embedding process is followed mathematically as shown in the figure 3.

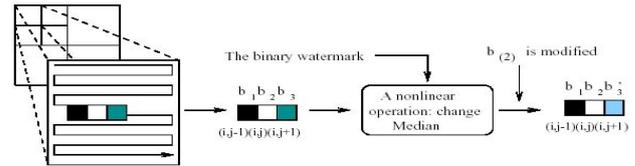


Fig. 3: Xie Embedding Watermark Method

Scheme of the proposed approach is showcased in figure.4

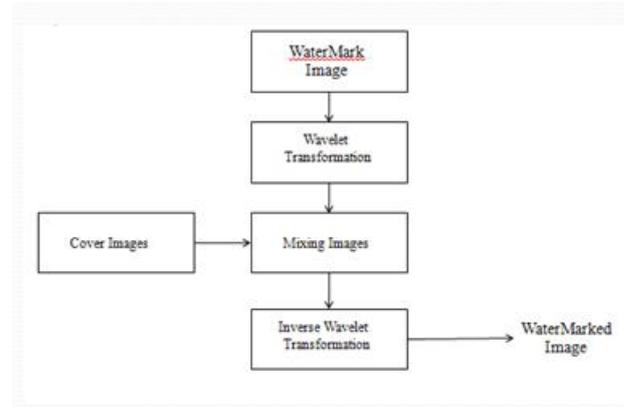


Fig. 4 Embedding Watermarking in image

3.2 The Watermark Embedding Phase

Basic DWT is implemented to embed the water mark in the image. Following are the steps in Embedding Algorithm:

Input: The color image $H(N \times N)$ and watermark $W_p(2M \times 2M)$

Output: The embedded image (watermarked image)

Step 1: Applying one level DWT on the image H . to get the blocks

Step 2: Applying three level on watermark W_p to obtain four blocks. W_p^1 (LL band), W_p^2 (LH band), W_p^3 (HL band), W_p^4 (HH band)

Step 3: Select the blocks $W_p^1, W_p^2, W_p^3, W_p^4$ of watermark and transform to binary streams. $WR^n(K), WG^n(K)$ and $WB^n(K)$ respectively where $K = 1, 2, \dots, N/8 \times N/8$

Step 4: For $n = 1$ to 4, do steps 1 to 3. Transform the blocks of original image to binary bit streams. For $K = 1$ to $M \times M$, do step as follows. If K is odd, embed watermark into original image according to the logic rules

Step 5: Applying one level IDWT to obtain watermarked image.

3.3 Attacks possible on embedded image

This section demonstrates the potentiality of embedding algorithm implemented in the work and focus on robustness and reliability. The figure 5 shows the possible assumed attacks in the present study.

- i. **Active attacks:** Here, the hacker tries deliberately to eradicate the watermark or simply make it undetectable. This is a big issue in copyright protection, fingerprinting or copy control.
- ii. **Passive attacks:** In this case, the attacker would try to remove the watermark but simply attempting to determine if a given mark is present or not. As the reader should understand, protection against passive attacks is of the utmost importance in covert communications where the simple knowledge of the presence of watermark itself poses immense vulnerability in the future.

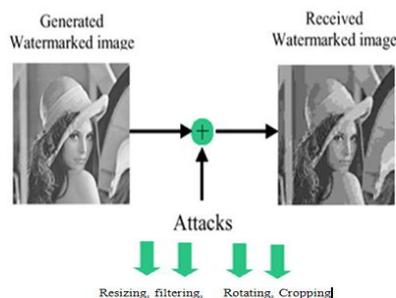


Fig. 5 shows the various attacks possible on the embedded water image

3.4 Image operations used in the Implementation: Image resizing, Image cropping and Image filtering.

- i. **Image resizing:** Aim of this attack on watermarked image is either to remove or to identify the watermark. Averaging every pixel in certain number leads to reduction dimensionality of image [15]. In this section it is to attempt the technique for removing or tracing the watermark of original image. The attack would be performed with the following steps.

Input: Watermarked Image

Output: Attacked watermarked image

Step 1: Select the watermarked image $I_w (N \times N)$

Step 2: Changing the pixel format

Step 3: Transform to the data type of the pixels to double and converting to single row of matrix.

Step 4: Calculating the average of certain number of pixels in overlapped manner

Step 5: Displaying the attacked output image.

- ii. **Image Cropping Attack:** Using the crop tool of image editing program which would draw a box around a selected portion of your digital image. When you execute the crop tool the remaining picture contains only what was inside the drawn box. A standard crop tool allows you to draw a rectangle of any height and width ratio. Most image editing programs would allow you to crop to a constrained ratio or proportion like 8 by 10 or 5 by 7 [14]. Some image editing programs will allow you to crop with a constrained ratio to resize and scale the image to a required print resolution in one

operation.

Input : Watermarked Image

Output: Cropped Watermarked Image

Step 1: Selecting the watermarked image $I_w (N \times N)$

Step 2: Change the pixel format of the image into double.

Step 3: Transforming the pixels to single row of matrix.

Step 4: Applying the Cropping Operation on certain selected portion of image “Crop (I_w)”

Step 5: Displaying the output attacked image.

iii. Image Filtering Attack

Input: Watermarked Image

Output: Attacked watermarked image

Step 1: Select the watermarked image $I_w (N \times N)$

Step 2: Changing the pixel format

Step 3: Transform to the data type of the pixels to double and converting to single row of matrix

Step 4: Applying the Filter operation function on both transformed pixel format and pixel in double format

Step 5: Displaying the attacked output image on filtering operation.

3.5 Watermark Extraction Algorithm

Input: Colour embedded image (Watermarked Image) I_m^n

$(N \times N)$ Output: The retrieved watermark image $R_i (2M \times 2M)$

Step 1: Apply three level DWT on R, G and B planes of embedded image I_m^n

Step 2: Select the blocks of $I_m^1, I_m^2, I_m^3, I_m^4$ and transform each block to binary streams

Step 3: Select original image and applying DWT on original image & transform the blocks to binary streams

Step 4: Subtract the bit streams of watermarked and original image

Step 5: Apply IDWT on difference to obtain the coefficients

Step 6: Displaying the extracted watermark.

3.6 De-blurring with Cropping with Normal, Wiener filter and DCT compression Technique

The most important technique for removal of blur in images due to linear motion or unfocused optics is the Wiener filter. From a signal processing, blurring due to linear motion in a photograph is the result of poor sampling [16]. Each pixel in a digital representation of the photograph should signify the intensity of a single stationary point in front of the camera [6]. Unfortunately, if the shutter speed is too sluggish and the camera is in motion, a given pixel will be a mix of intensities from points along the line of the camera's motion. $I = \text{imcrop}$ creates an interactive crop image and the image displayed in the current figure is called the target image. The crop image tool is a moveable, resizable rectangle that we can position interactively using the mouse. When the crop image tool is active, the pointer changes to cross hairs when you budge it over the target image [17]. Using the mouse, we can specify the crop rectangle by clicking and dragging the mouse.

IV. RESULTS OF THE EXPERIMENT

To measure the feasibility of the proposed scheme, we have conducted a series of experiments. The color host images include “Academy building” and “Logo”. The watermark is a visually recognizable binary image with the size of 88 X 88 pixels. The watermark size is 88 X 88, which is calculated according to the selected embedded locations and the PSNR value. The blocks size of the I^3 , I^4 , I^7 and I^8 (refer to figure 6 - 15.) from the R and B bands for embedding are 32768. Because the secret sharing scheme will cause the size of the watermark four times larger. The watermark is decided to 88 x 88.

The image processing operations that have been applied are cropping, blurring, sharpening, scaling, JPEG compression, brightness adjustment and contrast adjustment[5]. The experimental results are listed one by one in the following sections. In the analysis of various images PSNR is

calculated $PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right)$ **4.1 Embedding**

Watermark

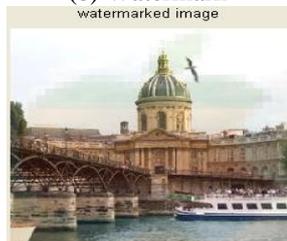
The output of embedded watermark is shown in the following image which shows separately original, watermark and embedded image.



(a) Image



(b) Watermark



(c) Embedded Watermarked Image

Fig. 6: (a,b,c)

The Embedding Algorithm Simulation is carried out in the MATLAB environment and the output images are shown in the figures.

4.2 Attacks

The selection reveals robustness and reliability of embedding algorithm used in this project. The various attacks on watermarked image are attempted and PSNR value calculated corresponding to the all mechanisms of attacks.

i. Attack of watermarked image with filtering

This attack involves both normal and Weiner filtering techniques. The simulation of this technique gives the output image as follows.

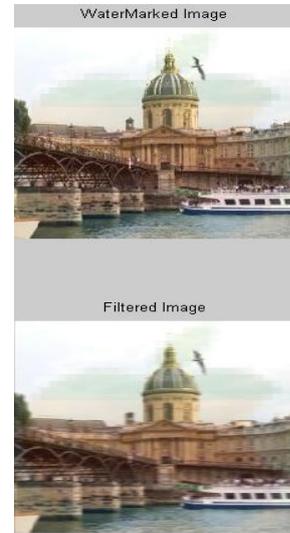


Fig. 7: Filtered Image

The PSNR value above attacked image is found to be equal to 0.041538.

ii. Attack of watermarked image resizing

The image resizing algorithm as discussed in the preceding section and produces the output image as follows.



Fig. 8: Resized Image

The PNSR value for above image is equal to 0.09238.

iii. The attack of watermarked image cropping

The cropping image algorithm is implemented that simulates the following images as output.



Fig. 9: cropped Image

The PSNR value is 0.0374

4.3 Watermark Extraction and De-Blurring

The watermark has been extracted as discussed in earlier section. The extracted watermark is noisy and blurred. The PSNR value of this image is calculated and is found to be equal to 25.5203. The simulation of watermark detection algorithm produces the following output.



Fig. 10: Extracted watermark

The noisy extracted watermark is developed with various de-blurring techniques. And therefore PSNR can be enhanced that leads to clarity of the image. In this paper various de-blurring algorithms are attempted for filtering the redundant information. The foregoing analysis shows the simulation of filtering techniques to de-blur the noisy extracted watermark. The detailed description and functionality of the de-blurring algorithms implemented in the paper are already given in the preceding chapter. The de-blurring techniques play a vital role in watermarking task.

i. Normal and Wiener Filer

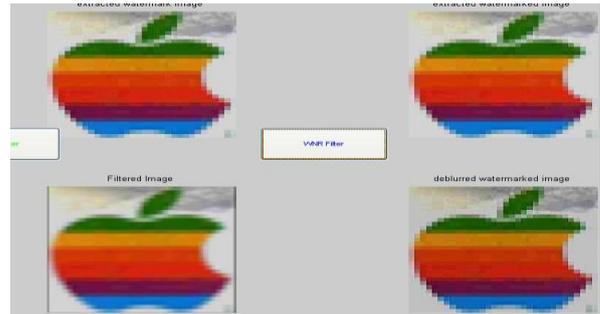


Fig. 11: Normal and Wiener Filtered Watermark

PSNR for normal filter 29.59 and Wiener filter 31.92

ii. De-Blurring with DCT Compression



Fig. 12: DCT Compressed image

PSNR value is 38.7791

iii. De-blurring with Cropping technique

Cropping removes some parts of an image. The experiment crops the three host images with different areas. According to the experiment, when the remaining cropped areas is 448 x 448, the accuracy rate of the watermark with the correction is nearly 90%, which can be identified by human eyes.



Fig. 13: Image cropped

iv. Cropping with Noise treatment

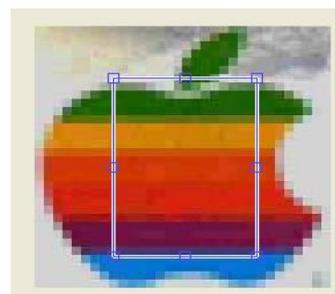


Fig. 14: Image cropped with noise treatment

In this paper, a copyright protection scheme for color images using discrete wavelet transform (DWT) is demonstrated. The implementation verifies robustness and reliability against various attacks on the embedded watermarked image. This scheme is suitable for color images. And it is also noted that PSNR in the de-blurred images enhances successively with implementation of normal, Wiener filtering, DCT and JPEG Compression techniques.

Furthermore, the advantages of the previously proposed scheme are still preserved in the improved proposed algorithm.

(1) It does not modify the host image, and therefore is suitable for unchangeable images,

(2) It is secure because of the employment of watermark.

(3) It is robust according to the experimental results with the calculated parameter of PSNR.

V. CONCLUSION

The investigation has shown that proposed technique is of an immense potentiality in secure transmission of data in the networks. The processes of embedding, extracting and de-blurring have been successfully simulated to reveal mechanism of avoiding the theft of secret watermark for sustained authenticity of the owner. The noisy retrieved image is de-blurred by employing various algorithms given in the paper and corresponding SNR is also estimated. An interesting point noted in the attempt is that performance of all proposed algorithms is well appreciated. Reliability and robustness of embedding algorithm have been examined carefully in the way to the analyze in case of image information embedded is being retrieved by any third party as a opponent.

References

- [1] Abdelwahab, A.A. & Hassan, L.A., 2008. A discrete Wavelet transform based technique for image data hiding. In *Proceedings of 25th National Radio Science Conference, NRSC 2008*. Egypt, 2008. March 18-20.pp.1-9.
- [2] Areepongsa, S., Kaewkamnerd, N., Syed, Y.F. & Rao, K.R., 2000. Exploring on steganography for low bit rate wavelet based coder in image retrieval system. In *Proceedings of IEEE TENCON*. Kuala Lumpur, Malaysia, 2000.
- [3] Ashtiyani, M., Birgani, P.M. & Hosseini, H.M., 2008. Chaos-based medical image encryption using symmetric cryptography. In *Proceedings of IEEE 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA 2008)*, 2008. 7-11 April. pp.1-5.
- [4] Cancelli, G., Doërr, G.J., Barni, M. & Cox, I.J., 2008. A comparative study of +/-1 steganalyzers. In *Proceedings of IEEE 10th Workshop on Multimedia Signal Processing*. Queensland, Australia, 2008. 8-10 Oct. pp.791-796.
- [5] Chang, C.C., Chen, T.S. & Chung, L.Z., 2002. A steganographic method based upon JPEG and quantization table modification. *Information Sciences*, 141(1-2), pp.123-38.
- [6] Deguillaume, F., Voloshynovskiy, S. & Pun, T., 2003. Secure hybrid robust watermarking resistant against tampering and copy attack. *Signal Processing*, 83(10), pp.2133-70.
- [7] Delachaye, J.P., 1996. Information noyée, information cache. *Pour la Science*, 229, pp.142-46. www.apprendre-en-ligne.net/crypto/stegano/229_142_146.pdf.
- [8] Denis, T.S., 2006. Cryptography for Developers. In *Syngress*, 2006.
- [9] Jarvis, J.F., Judice, C.N. & Ninke, W.H., 1976. A Survey of Techniques for the Display of Continuous-tone Pictures on Bilevel Displays. *Computer Graphics and Image Processing*, 5(1), pp.13-40.
- [10] Nirinjan, U.C. & Anand, D., 1998. Watermarking medical images with patient information. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Hong Kong, China, 1998. 29 Oct-1 Nov. pp.703-706.
- [11] Saenz, M., Oktem, R., Egiazarian, K. & Delp, E., 2000. Color image wavelet compression using vector morphology. In *Proceedings of the European Signal Processing Conference*. Tampere, Finland, 2000. 5-8 September. pp.5-8.
- [12] Van Der Weken, D., Nachtegael, M. & Kerre, E., 2004. Using similarity measures and homogeneity for the comparison of images. *Image and Vision Computing*, 22(9), pp.695-702.
- [14] Wang, K.W., 2009. Image encryption using chaotic maps. In Kocarev, L., Galias, Z. & Lian, S. *Intelligent computing based on chaos*. Springer. p.345.
- [15] Wang, Y. et al., 2007. image encryption method based on chaotic map. In *Proceedings of IEEE 2nd Conference on Industrial Electronics and Applications (ICIEA)*. harbin, China, 2007. 23-25 May. pp.2558-2560.
- [16] Yu, Y.H., Chang, C.C. & Lin, I.C., 2007. A new steganographic method for color and grayscale image hiding. *Computer Vision and Image Understanding*, 107(3), pp.183-94.
- [17] Zeghid, M. et al., 2006. A modified AES based algorithm for image encryption. *International Journal of Computer Science and Engineering*, 1(1), pp.70-75

BIOGRAPHIES



G. Srinivas Reddy, received M.Sc. in Applied Electronics from Osmania University, Hyderabad, India, holds an M.Tech in Computer Science and Engineering from Jawaharlal Nehru Technological University, Hyderabad, A.P., India. He has 9 years of experience in teaching at various Engineering Colleges. He is presently Assistant Professor at Mahatma Gandhi Institute of Technology [MGIT], Hyderabad, A.P, INDIA. His areas of research include Evolutionary algorithms, Artificial Neural Networks, Computational physics, Digital Image Processing, Network Security and other Emerging areas of computing sciences. He is pursuing Ph.D. He can be reached at



T.Venkat Narayana Rao, received B.E in Computer Technology and Engineering from Nagpur University, Nagpur, India, M.B.A (Systems), holds a M.Tech in Computer Science from Jawaharlal Nehru Technological University, Hyderabad, A.P., India and a Research Scholar in JNTU, Kakinada. He has 21 years of vast experience in Computer Science and Engineering areas pertaining to academics and industry related I.T issues. He is presently working at , Department of Computer Science and Engineering, Hyderabad Institute of Technology and Management [HITAM], Gowdavally, R.R.Dist., A.P, INDIA. He is currently working on research areas which include Digital Image Processing, Digital Watermarking, and other Emerging areas of Information Technology. He can be reached at



Dr.A.Govardhan, BE in computer Science and Engineering from Osmania University College of Engineering, Hyderabad, M.Tech from Jawaharlal Nehru University, Delhi and Ph.D from Jawaharlal Nehru Technological University, Hyderabad. Worked as Principal and Professor, C.S.E, College of Engineering, Jawaharlal Nehru Technological University-Hyderabad, Nachupally Karimnagar, A P, India. Presently working as Director of Evaluation JNTU, Hyderabad. A member of Standing Committee for Academic Senate, JNT University Hyderabad and Academic Advisory Committee (AAC), UGC-Academic Staff College, JNT University Hyderabad. He is the Chairman for Post Graduate Board of Studies (BOS) in Computer Applications, Yogi Vemana University, and Kadapa. A member, BOS for CSE and IT (UG and PG), JNT University Hyderabad and VR Siddhartha

Engineering College, Vijayawada. He was the Chairman for Board of Studies (Computer Science and Engineering) during 2008-2009 for UG and PG at Dept. of CSE, JNTUH College of Engineering Hyderabad and Member BOS at School of Information Technology, JNTU Hyderabad. He was the Co-convener for EAMCET 2009. He has been a committee member for various International and National conferences including PAKDD2010, IKE10, ICETCSE-2010 ICACT-2008, NCAI06. He is a Coordinator for the ongoing Research Project on Telugu in IT. Government of A.P. He is also a member in various professional bodies including CSI, ISTE, IAENG, FSF and WASET. He has been listed as one among the Top Three Faculty in JNTU Hyderabad made by Outlook Survey for the year 2008. He has 135 research publications at International/National Journals and Conferences. He is Member, Editorial Board of many reputed International Journals of Emerging Technologies and Applications in Engineering Technologies. He has been a program committee member for various International and National conferences. He has delivered number of Keynote addresses and invited lectures. His areas of interest include Databases, Data Warehousing & Mining, Information Retrieval, Computer Networks, Image Processing and Object Oriented Technologies.

Static Response of a Involute Pairs Of Gears In Application Scenario (Validation Using AGMA Standards) By Using FEM Cutting Edge Technology

Vikas Kumar¹, Mohd. Parvez²

¹Rattan of Technology and Management, Saveli, Palwal,

²Al-Falah School of Engg. and Technology, Dhauj Faridabad,

Abstract: This paper investigates the characteristics of an involutes gear system including contact stresses, bending stresses, and the transmission errors of gears in mesh. Gearing is one of the most critical components in mechanical power transmission systems. Transmission error is considered to be one of the main contributors to noise and vibration in a gear set. Transmission error measurement has become popular as an area of research on gears and is possible method for quality control. To estimate transmission error in a gear system, the characteristics of involute spur gears were analyzed by using the finite element method. The contact stresses were examined using 2-D FEM models. The bending stresses in the tooth root were examined using a 3-DFEM model. Current methods of calculating gear contact stresses use Hertz's equations, which were originally derived for contact between two cylinders. To enable the investigation of contact problems with FEM, the stiffness relationship between the two contact areas is usually established through a spring placed between the two contacting areas. This can be achieved by inserting a contact element placed in between the two areas where contact occurs. The results of the two dimensional FEM analyses from ANSYS are presented. These stresses were compared with the theoretical values. Both results agree very well.

This indicates that the FEM model is accurate. The finite element method is very often used to analyze the stress state of an elastic body with complicated geometry, such as a gear. There have been numerous research studies in the area [1],[2]

I. INTRODUCTION

1.1 Background

When one investigates actual gears in service, the conditions of the surface and bending failure are two of the most important features to be considered. The finite element method is very often used to analyze the stress states of elastic bodies with complicated geometries, such as gears. There are published papers, which have calculated the elastic stress distributions in gears. In these works, various calculation methods for the analysis of elastic contact problems have been presented. The finite element method for two-dimensional analysis is used very often. It is essential to use a three-dimensional analysis if gear pairs are under partial and non uniform contact. However, in the three dimensional calculation, a problem is created due to the large computer memory space that is necessary. In this chapter to get the gear contact stress a 2-D model was used Because it is a nonlinear problem it is better to keep the number of nodes and elements a slow as possible. In the

bending stress analysis the 3-D model and 2-D models are used for simulation

1.2 Analytical Procedure

From the results obtained in earlier the present method is an effective and accurate method, which is proposed to estimate the tooth contact stresses of a gear pair. Special techniques of the finite element method were used to solve contact problems. Using the present method, the tooth contact stresses and the tooth deflections of a pair of spur gears analyzed by ANSYS 13. Since the present method is a general one, it is applicable to many types of gears. In early works, the following conditions were assumed in advance:

- There is no sliding in the contact zone between the two bodies

- The contact surface is continuous and smooth

Using the present method ANSYS can solve the contact problem and not be limited by the above two conditions. A two-dimensional and an asymmetric contact model were built. First, parameter definitions were given and then many points of thin involute profile of the pinion and gear were calculated to plot an involute profile using a cylindrical system. The equations of an involute curve below were taken from Buckingham [3]:

$$r = r_b * (1 + \beta^2)^{1/2} \quad 1.2.1$$

$$\psi = \theta + \pi / 2n_1 - \xi \quad 1.2.2$$

$$\theta = \tan \phi_1 - \phi_1 \quad 1.2.3$$

where r = radius to the involute form, $b r$ = radius of the base circle

$$\beta = \xi + \phi_1$$

θ = vectorial angle at the pitch circle

ξ = vectorial angle at the top of the tooth

ϕ_1 = pressure angle at the pitch circle

pressure angle at radius r

One spur tooth profile was created using above equations, shown in Figure 1, as are the outside diameter circle, the dedendum circle, and base circle of the gear. Secondly, in ANSYS from the tool bars using "CREATE", "COPY", "MOVE", and "MESH" and so on, any number of teeth can be created and then kept as the pair of gear teeth in contact along the line of the action. The contact conditions of gear teeth are sensitive to the geometry of the contacting surfaces, which means that the element near the contact zone needs to be refined. It is not recommended to have a fine mesh everywhere in the model, in order to reduce the computational requirements. There are

two ways to build the fine mesh near the contact surfaces. One is the same method as presented in later, a fine mesh of rectangular shapes were constructed only in the

contact areas. The other one, "SMART SIZE" in ANSYS, was chosen and the fine mesh near the contact area was automatically created. A FEM gear contact model was generated as shown in Figure .

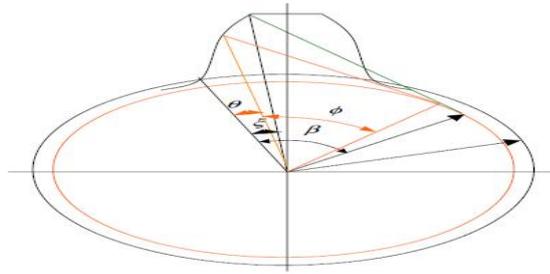


Figure 1 Involumentry of a spur gear

Thirdly, proper constraints on the nodes were given. The contact pair was inserted between the involute profiles, the external loads were applied on the model from ANSYS "SOLUTION > DEFINE LOAD > FORCE / MOMENT", and finally, ANSYS was run to get the solution

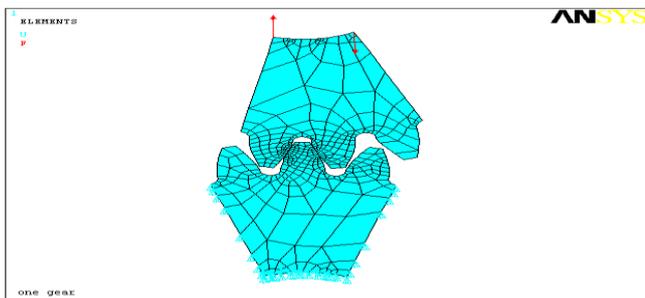


Figure 2 Gear contact stress model

II. Rotation Compatibility of the Gear Body

In order to know how much load is applied on the contact stress model and the bending stress model, evaluating load sharing between meshing gears is necessary. It is also an important concept for transmission error. It is a complex process when more than one-tooth pair is simultaneously in contact taking into account the composite tooth deflections due to bending, shearing and contact deformation. This section presents a general approach as to how the load is shared between the meshing teeth in spur gear pairs. When the gears are put into mesh, the line tangent to both base circles is defined as the line of action for involute gears. In one complete tooth mesh cycle, the contact starts at points A shown in Figure 3 where the outside diameter circle, the addendum circle of the gear intersects the line of action. The mesh cycle ends at point E as shown in Figure 4 where the outside diameter of the pinion intersects the line of action.

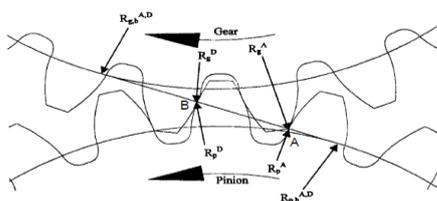


Figure 3 Illustration of one complete tooth meshing cycle

Consider two identical spur gears in mesh. When the first tooth pair is in contact at point A it is between the tooth tip of the output gear and the tooth root of the input gear (pinion). At the same time a second tooth pair is already in contact at point D in Figure 3. As the gear rotates, the point of contact will move along the line of action APE. When the first tooth pair reaches point B shown in Figure 4, the second tooth pair disengage at point E leaving only the first tooth pair in the single contact zone. After this time there is one pair of gear in contact until the third tooth pair achieves in contact at point A again. When this tooth pair rotates to point D, the another tooth pair begins engagement at point A which starts another mesh cycle.

After this time there are two pairs of gear in contact until the first tooth pair disengage at point E. Finally, one complete tooth meshing cycle is completed when this tooth pair rotates to point E. To simplify the complexity of the problem, the load sharing compatibility condition is based on the assumption that the sum of the torque contributions of each meshing tooth pair must equal the total applied torque.

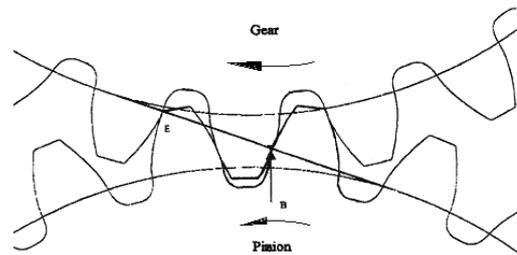


Figure 4 Different positions for one complete tooth meshing cycle

Analytical equations can also be developed for the rotation of the gear and pinion hubs, including the effects of tooth bending deflection and shearing displacement and contact deformation . In the pinion reference frame, it is assumed that the pinion hub remains stationary, while the gear rotates due to an applied torque. Considering the single pair contact zone at point B, the condition of angular rotation of the gear body will then be given by

$$\text{for the pinion} \quad \theta_p^B = (B_p^B + H_p^B) / R_p^B \quad 2.1$$

$$\text{and for the gear} \quad \theta_g^B = (B_g^B + H_g^B) / R_g^B \quad 2.2$$

where B_p^B and B_g^B are the tooth displacement vectors caused by bending and shearing for pairs B of the pinion and gear respectively, H_p^B and H_g^B are the contact deformation vectors of tooth pair B of the pinion and gear respectively. θ_p^B denotes the transverse plane angular rotation of the pinion body caused by bending deflection, shearing displacement and contact deformation of the tooth pair B while the gear is stationary. Conversely, for the gear rotation while the pinion is stationary, θ_g^B gives the transverse plane angular rotations of the gear body.

III. Gear Contact Stress

One of the predominant modes of gear tooth failure is pitting. Pitting is a surface fatigue failure due to many repetitions of high contact stress occurring on the gear tooth surface while a pair of teeth is transmitting power. In other words, contact stress exceeding surface endurance strength

with no endurance limits or a finite life causes this kind of failure. The AGMA has prediction methods in common use. Contact failure in gears is currently predicted by comparing the calculated Hertz stress to experimentally determined allowable values for the given material. The details of the subsurface stress field usually are ignored. This approach is used because the contact stress field is complex and its interaction with subsurface discontinuities are difficult to predict. However, all of this information can be obtained from the ANSYS model. Since a spur gear can be considered as a two-dimensional component, without loss of generality, a plane strain analysis can be used. The nodes in the model were used for the analysis. The nodes on the bottom surface of the gear were fixed. A total load is applied on the model. It was assumed to act on the two points shown in Figure 1 and three points in Figure 4.

There are two ways to get the contact stress from ANSYS. Figure 4 shows the first one, which is the same method as we use earlier to create the contact element COCNTA 48 and the rectangular shape fine mesh beneath the contact surfaces between the contact areas. Figure 5 shows the enlarged-area with a fine mesh which is composed of rectangular shapes.

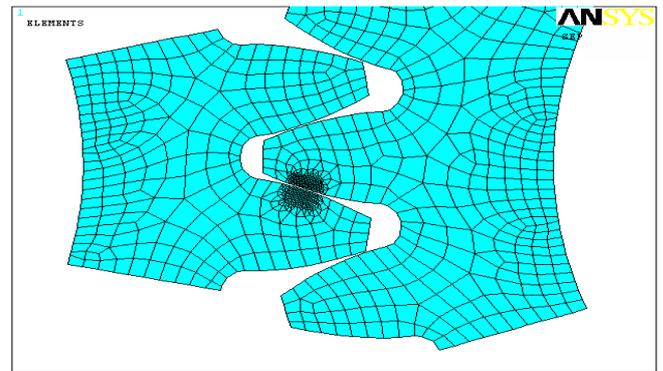


Figure 8 A fine mesh near contact areas

Fig. 6 shows the normal contact stress along the contact areas. The results are very similar to the results in the two cylinders in our solution. Fig. 7 presents how to mesh using a second method. Different methods should show the close results of maximum contact stress if the same dimension of model and the same external loads are applied on the model. If there is a small difference it is likely because of the different mesh patterns and restricted conditions in the finite element analysis and the assumed distribution form of the contact stresses in the contact zone.



Figure 5 FEM Model of the gear tooth pair in contact

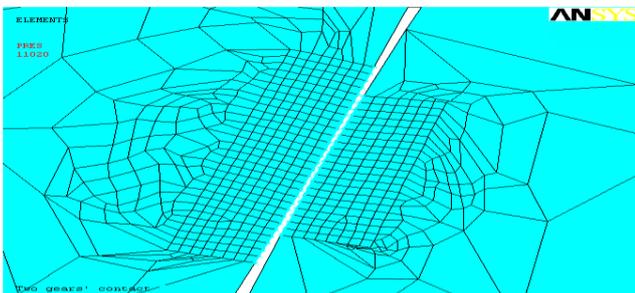


Figure 6 Fine meshing of contact areas

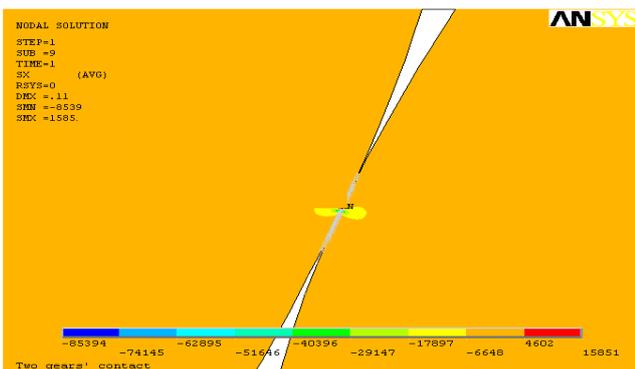


Figure 7 Contact stress along contact areas

IV. The Lewis Formula

There are several failure mechanisms for spur gears. Bending failure and pitting of the teeth are the two main failure modes in a transmission gearbox. Pitting of the teeth is usually called a surface failure. This was already discussed in the last section. The bending stresses in a spur gear are another interesting problem. When loads are too large, bending failure will occur. Bending failure in gears is predicted by comparing the calculated bending stress to experimentally-determined allowable fatigue values for the given material. This bending stress equation was derived from the Lewis formula. Wilfred Lewis (1892) [4] was the first person to give the formula for bending stress in gear teeth using the bending of a cantilevered beam to simulate stresses acting on a gear tooth shown in Figure 9 are Cross-section = $b * t$, length = 1, load = $t F$, uniform across the face. For a rectangular section, the area moment of inertia is

$$I = bh^3/12 \tag{4.1}$$

$M = F_t l$ and $c = t/2$, stress then is

$$\sigma = \frac{M}{I/c} = \frac{F_t l (t/2)}{bt^3/12} = \frac{6F_t l}{bt^2} \tag{4.2}$$

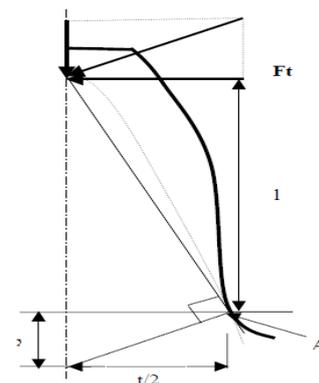


Figure 9 Length dimension used in determining bending tooth stress

Where b = the face width of the gear. For a gear tooth, the maximum stress is expected at point A, which is a tangential point where the parabola curve is tangent to the curve of the tooth root fillet called parabola tangential method. Two points can be found at each side of the tooth root fillet. The stress on the area connecting those two points is thought to be the worst case. The crack will likely start from the point A.

From similar triangles $\tan\alpha = \frac{t/2}{l} = \frac{1}{4x}$ where $l = t^2/4x$
 4.3

Substituting (4.7) into (4.6):

$$\sigma = \frac{F_t}{bY} \frac{t^2/4x}{t^2/4x} = \frac{3F_t}{2bY} = \frac{3F_t P_d}{2bY} = F_t P_d \frac{3}{2bY} = 4.4$$

where P_d = diametral pitch

$$Y = 2xP_d = \text{Lewis form factor} \quad 4.5$$

$$\sigma_t = \frac{F_t P_d K_a K_s K_m}{b Y_j K_v} \quad 4.6$$

where K_a = application factor, K_s = size factor,

K_m = load distribution factor, K_v = dynamic factor,

F_t = normal tangential load, Y_j = Geometry factor.

V. The Two Dimensional Model.

Fatigue or yielding of a gear tooth due to excessive bending stresses are two important gear design considerations. In order to predict fatigue and yielding, the maximum stresses on the tensile and compressive sides of the tooth, respectively, are required. In the past, the bending stress sensitivity of a gear tooth has been calculated using photo elasticity or relatively coarse FEM meshes. However, with present computer

driving gear, the iterative procedures were used to solve the static equilibrium of the gear pair and to calculate the load distribution on the contact lines and the static transmission error. However the contact deformation was excluded in those models. This section considers a FEA model, which was used to predict static transmission error of a pair of spur gears in mesh including the contact deformation. The model involves the use of 2-D elements, coupled with contact elements near the points of contact between the meshing teeth. When one pair of teeth is meshing one set of contact elements was established between the two contact surfaces, while when two pairs of teeth are meshing two sets of contact elements were established between the two contact bodies. When gears are unloaded, a pinion and gear with perfect involutes profiles should theoretically run with zero transmission error. However, when gears with involute profiles are loaded, the individual torsional mesh stiffness of each gear changes throughout the mesh cycle, causing variations in angular rotation of the gear body and subsequent transmission error. The theoretical changes in the torsional mesh stiffness throughout the mesh cycle match the developed static transmission error using finite element analysis shown in Figure 8.

The literature available on the contact stress problems is extensive. But that available on the gear tooth contact stress problem is small, especially for transmission error including the contact problem. Klenz [5] examined the spur gear contact and bending stresses using two dimensional FEM. Coy and Chao [6] studied the effect of the finite element grid size on Hertzian deflection in order to obtain the optimum aspect ratio

at the loading point for the finite element grid. Gatcombe and Prowell [7] studied the Hertzian contact stresses and duration of contact for a very specific case, namely a particular rocket motor gear tooth. T say [8] has studied the bending and contact stresses in helical gears using the finite element method with the tooth contact analysis technique. However, the details of the techniques used to evaluate the transmission error including contact stresses were not presented.

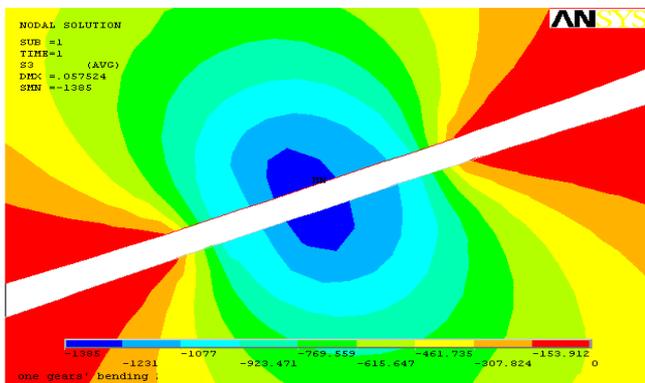


Figure 10 The distribution of contact stresses between two teeth

6 The Transmission Error. The static transmission error is expressed as a linear displacement at the pitch point. A kinematic analysis of the gear mesh allows determining the location of contact line for each loaded tooth pair. These contact lines were discretized. The total length of lines of contact grows with the applied load. For each position of the

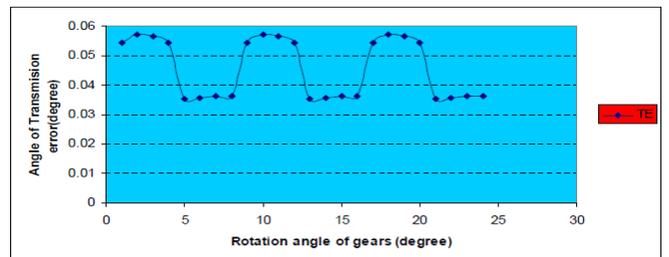


Figure 11 static transmission error from ansys

VI. Conclusion

Mesh stiffness variation as the number of teeth in contact changes is a primary cause of excitation of gear vibration and noise. This excitation exists even when the gears are perfectly machined and assembled. Numerical methods using 2-D FEM modelling of toothed bodies including contact elements have been developed to analyze the main static transmission error for spur gear pairs. Numerous simulations allowed validating this method and showed that a correct prediction of transmission error needed an accurate modelling of the whole toothed bodies.

The elasticity of those bodies modifies the contact between loaded tooth pairs and the transmission error variations. The developed numerical method allows one to optimize the static transmission error characteristics by introducing the suitable tooth modifications. These offer interesting possibilities as first steps of the development of a transmission system and can be also successfully used to improve to control the noise and vibration generated in the transmission system.

References

- [1] Umezawa, K., 1988, "Recent Trends in Gearing Technology", JSME International Journal Series III Vol.31, No. 2, pp 357-362
- [2] Smith, J. O. Liu, C. K., "Stresses Due to Tangential and Normal Loads on an Elastic Solid with Applications to Some Contact Stress Problems", *Journal of Applied Mechanics*
- [3] Buckingham, E., 1949, "Analytical Mechanics of Gears", McGraw-Hill, New York
- [4] Hamrock, B. J., Jacobson, S. R., "Fundamentals of Machine Elements".
- [5] Klenz, S. R., 1999, "Finite Element Analyses of A Spur Gear Set", M.Sc. Thesis, Dept. of Mechanical Engineering, University of Saskatchewan.
- [6] Coy, J. J., Chao, C. H. S., 1982, "A method of selecting grid size to account for Hertz deformation in finite element analysis of spur gears", *Trans. ASME, J. Mech. Design* 104 759-766
- [7] Gatcombe, E.K., Prowell, R.W., 1960, "Rocket motor gear tooth analysis", (Hertzian contact stresses and times) *Trans. ASMA, J. Engng Industry*
- [8] Tsay, C.B., 1988, "Helical gears with involute shaped teeth: Geometry, computer simulation, tooth contact analysis, and stress analysis", *Trans, J. Mechanisms, Transmissions, and Automation in Design*

Thermal Stress Analysis of Orthotropic Graded Rotating Discs

Gaurav Kansal¹ Mohd. Parvez²

¹BS Anangpuria Institute of Technology and Management, Alampur, Faridabad,
²Al-Falah School of Engg. and Technology, Dhauj Faridabad,

Abstract: The present study deals with stress analysis on orthotropic graded rotating annular discs subjected to temperature distributions parabolically decreasing with radius. I have used infinitesimal deformation theory of elasticity and for graded parameters power law functions in the solution procedure. With the increasing temperature, the tangential stress component decreased at the inner surface whereas increased at the outer surface, and the radial stress component reduced gradually for all the temperature distributions. The magnitude of the tangential stress component was higher than ones of the radial stress component under the room temperatures for both discs. But, the tangential stress component decreased more at the inner surface whereas it increased at the outer surface when the temperature increased further. Finally, the radial displacement at the outer surface had higher value than that of the inner surface with the increasing temperature.

keyword: ANSYS, FGM, FGD,

I. Introduction

Composites are gradually being used as structural materials in many aerospace and automobile applications. The reinforcement in these composites is generally distributed uniformly. Functionally graded composite materials (FGMs) have been the subject of intense researches and attracted considerable attention in recent years. FGMs are being used as interfacial zone to improve the bonding strength of layered composites, to reduce the residual and thermal stresses in bonded dissimilar materials and as wear resistant layers in machine and engine components (Pindera et al., 1995; Erdoğan, 1995). One of the advantages of FGMs over laminates is that, due to continuous material property variation, there is no stress build-up at sharp material boundaries thus eliminating potential structural integrity issues such as delamination. Analysis of rotating discs is an important issue in mechanics and engineering applications.

An analytical solution for the stress analysis in the isotropic rotating disc or a disc under pressure can be found in literature (Timoshenko and Goodier, 1951). Çallioğlu et al. (2006) investigated elastic-plastic stress analysis of the curvilinearly rotating discs.

Sayman (2006) studied elastic-plastic stress analysis of a thermoplastic composite disc under uniform temperature distribution analytically and using a finite element commercial code (ANSYS).

Singh and Ray (2002) investigated creep in orthotropic aluminum-silicon carbide composite rotating disc by using Hill's anisotropic yield. In that study, the results obtained have been compared with the results obtained using von Mises yield criterion for the isotropic

composites. In all of these studies, tangential and radial elasticity moduli are constant, that is, the orthotropy degree is a constant. Since the mathematical problems arising are complicated, much of the work on FGMs has been carried out numerically. Nevertheless, the mechanical and mathematical modeling of FGMs is currently an active research area.

Horgan and Chan (1999a, b) investigated the stress response in rotating disks and pressurized hollow cylinder or disk made of functionally graded isotropic linearly elastic materials. They investigated a body with Young's modulus varying radially only. Horgan and Baxter (1996) examined the externally pressurized hollow **FGD**, stresses in functionally graded discs.

thermal stresses in the basic structural components of FGMs.

Jahed H, Dubey RN [1] has presented a modification of the Tsai-Wu criterion, needed in the case of the multi-criterion optimal design of thin-walled composite structure and a proposal of the evaluation of the load carrying capacity of multi-layered composites with respect to their failure mode.

Jahed H, Sherkatti S., [2] have presented a semi-analytical three-dimensional elasticity solution for rotating functionally graded disks. Their solution includes the responses of both of the hollow and solid disks and is a generalization of the two-dimensional plane-stress solution.

variation along the radius.

Leopold WR [5] has examined the stress analysis on orthotropic rotating annular disks subjected to various temperature distributions, such as uniform, linearly increasing and decreasing with radius temperatures.

Hosseini Kordkheili and Naghdabadi [6] have presented a semi-analytical thermoelasticity solution for hollow and solid rotating axisymmetric disks made of functionally graded cylinder (or disk) with radially orthotropic material.

Durodola and Attia (2000) investigated deformation and stresses in functionally graded rotating disks. They compared two methods, finite element method (ABAQUS) and direct numerical integration of governing differential equations, with each other.

Zenkour (2007) dealt with a solution for a rotating annular disk which is assumed to be graded in the radial direction according to a simple exponential-law distribution.

Chen et al. (2007) presented three-dimensional

analytical solution for a rotating disc of functionally graded materials with transverse isotropic. A significant amount of these studies has been done in order to see the effects of the FGMs (variation of the elasticity modulus only) on the isotropic discs and cylinders.

You et al. (2007) investigated stress analysis on the FG rotating circular discs under uniform temperature.

Zhang et al. (2003) provided an exact thermal stress solution for a functionally graded plate that has a circular hole, with the material properties and applied temperature varying arbitrarily in the radial direction.

Mohammadi and Dryden (2008) examined the role of nonhomogeneous stiffness on the thermoelastic stress field. In their contribution, Young's modulus and thermal expansion were considered to change symmetrically across the radius representing coatings on inner and outer radii of the beam.

Çallioğlu (2008) studied the stress analysis of the rotating hollow discs made of functionally graded materials under internal and external pressures. In that study, elasticity modulus and density change in the radial direction. In the study it is assumed that the isotropic material is of radially varying elasticity modulus E , density ρ and thermal expansion coefficient α and Poisson's ratio ν as a constant.

Closed-form solutions for stresses and displacements in functionally graded annular discs rotating at a constant angular velocity and subjected to temperature varying parabolically along the radial direction is obtained using the infinitesimal theory of elasticity and for functional graded case power-law function.

II. THERMAL STRESS ANALYSIS

Due to the fact that this problem is axisymmetric, the equilibrium equation for a rotating thin disc is:

$$\frac{rd\sigma_r}{dr} + \sigma_r - \sigma_\theta + \rho(r)\omega^2 r^2 = 0 \quad (2.1)$$

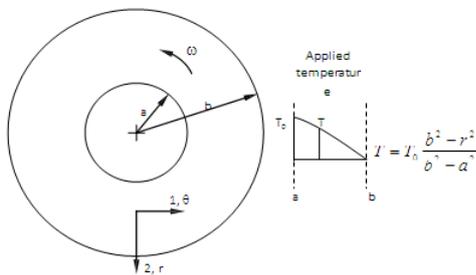


Figure 1. a functionally graded rotating disk under parabolic temperature distribution

Where, σ_r , σ_θ , ω and $\rho(r)$ are respectively, the radial stress, tangential stress, angular velocity and the radially varying material mass density. r is the radial distance, $r \neq 0$ and $a < r < b$. Here a and b are inner and outer radii of disc illustrated in Figure 1. The solution can be efficiently handled by using a special stress function that automatically satisfies the equilibrium in Equation 2.1. The particular stress-stress function relation with this property is given by

$$\sigma_r = \frac{F}{r}, \quad \sigma_\theta = \frac{dF}{dr} + \rho(r)\omega^2 r^2 \quad (2.2)$$

Where, $F = F(r)$ is the stress function.

The governing equation for the stress function is determined from the compatibility statement. For this axisymmetric case, the displacement field is of the form $u = u_r = u_r(r)$ and $u_\theta = 0$. Therefore, the strain field is given by:

$$\epsilon_r = \frac{du}{dr}, \quad \epsilon_\theta = \frac{u}{r}, \quad \epsilon_{r\theta} = 0 \quad (2.3)$$

Where, ϵ_r , ϵ_θ and u are the strains in radial and tangential directions and displacement component in the radial direction. Eliminating u from these equations develops the simple compatibility statement:

$$\epsilon_r = \epsilon_\theta + r \frac{d\epsilon_\theta}{dr} \quad (2.4)$$

Using Hooke's law for plane stress case, the strains are given by:

$$\epsilon_r = \frac{1}{E(r)} (\sigma_r - \nu\sigma_\theta) + \alpha(r)T(r) \quad (2.5)$$

$$\epsilon_\theta = \frac{1}{E(r)} (\sigma_\theta - \nu\sigma_r) + \alpha(r)T(r)$$

Where, E , α and T are respectively, elasticity modulus, thermal expansion coefficient and temperature change, and it is assumed that the disc is of material properties (E, ρ and α) and thermal change (T) varying through the radial section. Poisson's ratio ν is assumed that a constant due to the fact that its variation has much less practical significance than that in the other material properties. Using this result in the compatibility relation to, Equation 2.4 generates the governing equation:

$$r^2 F'' + rF' \left(1 - r \frac{E'(r)}{E(r)}\right) + F \left(\nu r \frac{E'(r)}{E(r)} - 1\right) = -\rho(r)\omega^2 r^3 \left(3 + \nu - r \frac{E'(r)}{E(r)}\right) - \rho'(r)\omega^2 r^4 - E(r)r^2 (\alpha'(r)T(r) + \alpha(r)T'(r)) \quad (2.6)$$

The superscript "n" represents derivatives with respect to r . Let us assume for the sake of argument that:

$$E(r) = E \left(\frac{r}{b}\right)^{n_1} \quad (2.7)a$$

$$\rho(r) = \rho \left(\frac{r}{b}\right)^{n_2} \quad (2.7)b$$

$$\alpha(r) = \alpha \left(\frac{r}{b}\right)^{n_3} \quad (2.7)c$$

$$T(r) = T_0 \left(\frac{b^2 - r^2}{b^2 - a^2}\right) \quad (2.8)d$$

Where, n_1 , n_2 and n_3 are dimensionless arbitrary constants (gradient parameters) and T_0 is the initial temperature at the inner surface of the functionally graded disc. The differential Equation 2.6 reduces to:

$$r^2 F'' + rF' \left(-n_1\right) + F \left(n_1 - 1\right) = -\rho\omega^2 \frac{(3 + \nu - n_1 + n_2)}{b^{n_2}} r^{n_2+3} - \frac{E\alpha T_0 n_3}{b^{n_1+n_2-2}(b^2 - a^2)} r^{n_1+n_2+1} + \frac{E\alpha}{b^{n_1+1}} \quad (2.9)$$

The stress function F can be written as:

$$F = C_1 r^{(n_1+m)/2} + C_2 r^{(n_1-m)/2} + A r_2^{n_1+3} + B r_1^{n_1+3} + C r_1^{n_1+3} \quad (2.10)$$

where $m = \sqrt{(n_1^2 - 4\nu n_1 + 4)}$, C_1 and C_2 are the integration constant

$$A = - \frac{\rho \omega^2 (3 + \nu - n_1 + n_2)}{b^{n_2} (n_2^2 + 6n_2 - n_1 n_2 - 3n_1 + \nu n_1 + 8)}$$

$$B = \frac{E \alpha T_0 n_3}{b^{n_1 + n_2} (b^2 - a^2) (n_3^2 + 2n_3 - n_1 n_3 + n_1 + \nu n_1)}$$

$$C = \frac{E \alpha T_0 (n_3 + 2)}{b^{n_1 + n_2} (b^2 - a^2) (n_3^2 + 6n_3 - n_1 n_3 - 3n_1 + \nu n_1 + 8)}$$

As $n_1 = n_2 = n_3 = 0$ and $T_0 = 0$, B and C terms are equal to zero and, the disc becomes the isotropic rotating disc and the stress function F is

$$F = C_1 r + \frac{C_2}{r} - \frac{3 + \nu}{8} \rho \omega^2 r^3 \tag{2.11}$$

The stress components can be obtained from the stress function in Equation 9 as:

$$\sigma_r = C_1 r^{(n_1 + m - 2)/2} + C_2 r^{(n_1 - m - 2)/2} + A r^{n_2 + 3} + B r^{n_1 + n} + C r^{n_1 + n + 2} \tag{2.12a}$$

$$\sigma_\theta = \frac{(n_1 + m)}{2} C_1 r^{(n_1 + m - 2)/2} + \frac{(n_1 - m)}{2} C_2 r^{(n_1 - m - 2)/2} + (n_2 + 3) A r^{n_2 + 3} + (n_1 + n_3 + 1) B r^{n_1 + n} + (n_1 + n_3 + 3) C r^{n_1 + n + 2} + \rho(r) \omega^2 r^2 \tag{2.12b}$$

$$C_1 = \frac{D_2 b^{-\frac{n_1 + m + 2}{2}} - D_1 a^{-\frac{n_1 + m + 2}{2}}}{b^m - a^m}$$

$$C_2 = \frac{D_1 b^m a^{-\frac{n_1 + m + 2}{2}} - D_2 a^m b^{-\frac{n_1 + m + 2}{2}}}{b^m - a^m}$$

Where,

$$D_1 = -A a^{n_2 + 2} - B a^{n_1 + n_3} - C a^{n_1 + n_3 + 2}$$

$$D_2 = -A b^{n_2 + 2} - B b^{n_1 + n_3} - C b^{n_1 + n_3 + 2} \tag{2.13}$$

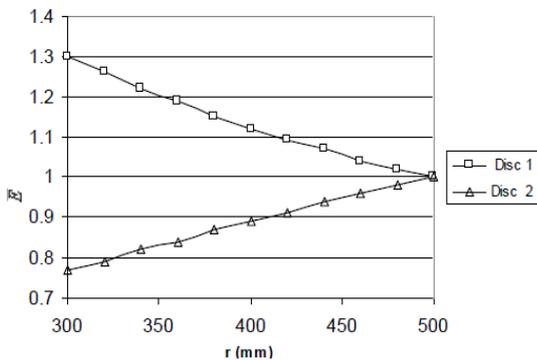


Figure 2. The normalized elasticity modulus distributions along the radial direction of discs.

Radial displacement component

Radial displacement by using the infinitesimal deformation theory of elasticity can be determined as:

$$u = \frac{r}{E(r)} (\sigma_\theta - \nu \sigma_r) + r \alpha(r) T(r) \tag{2.14}$$

III. Results and Discussion

In this paper, a thermal stress analysis is carried

out on functionally graded rotating annular discs by using an analytical solution including small deformation of theory of elasticity. The results are presented for Poisson's ratio $\nu = 0.3$ and angular velocity $\omega = 650 \text{ rad/s}$. The inner and outer radii of the discs are $a = 300 \text{ mm}$ and $b = 500 \text{ mm}$, respectively. Mechanical properties of the discs, such as elasticity modulus, density and thermal expansion coefficient, and temperature applied are assumed to be varying along the radial direction. The material coefficients are taken to be elasticity modulus $E = 150 \text{ GPa}$, density $\rho = 5600 \text{ kg/m}^3$ and thermal expansion coefficient $\alpha = 23 \times 10^{-6} \text{ 1/}^\circ\text{C}$, gradient parameters $n_1 = -0.5194$, $n_2 = -0.4873$ and $n_3 = 0.55236$ for Disc 1 and $n_1 = 0.5194$, $n_2 = 0.4873$ and $n_3 = -0.55236$ for Disc 2 (You et al., 2007). Temperature change is set to $T_0 = 0, 300$ and 600°C , respectively. If room temperature is considered as reference temperature, the room temperature should be added to the initial temperature, T_0 . Elasticity modulus, density and thermal expansion coefficient variations are given as normalized values along the radial direction of the discs in order to demonstrate the effects of FGMs on the discs. For E , ρ and α , the following formal normalized variables are used:

Figure 2 illustrates the variations of the normalized elasticity modulus E along the radial direction for Discs 1 and 2. As seen in this figure, the elasticity modulus is equal to elasticity modulus of the isotropic, homogeny disc at the outer surface. In the inner surface of the disc, elasticity modulus value increases gradually for Disc 1 while decreases for Disc 2. Figure 3 shows the variations of the normalized density ρ along the radial direction for Discs 1 and 2. As seen from this figure, the density value is equal to density of the isotropic, homogeny disc at the outer surface. Density value in the inner surface of the disc increases gradually for Disc 1 whilst decreases for Disc 2. Figure 4 depicts the variations of the normalized thermal expansion coefficient α along the radial direction for Discs 1 and 2. As seen from this figure, the thermal expansion coefficient value is equal to thermal expansion coefficient of the isotropic, homogeny disc at the outer surface. But, in the inner surface of the disc, contrary to E and ρ thermal expansion coefficient value increases gradually for Disc 2 whilst decreases for Disc 1. As seen from the last three figures, when the normalized elasticity modulus and density values for Disc 1 increase about 1.3 times of isotropic value, the normalized thermal expansion coefficient value decreases about 0.23 times of isotropic disc. These values for Disc 2 are the opposite of Disc 1 approximately. That is to say, thermal expansion coefficient decreases when elasticity modulus and density increase, or this situation is diametrically opposite. Variations of the temperature applied along the radial section of the discs are depicted in Figure 5. The temperature applied is of a variation decreasing parabolically from inner surface to outer surface along the radial direction. Figure 6 shows radial stress distributions along the radial section of both discs. Due to the boundary conditions, the radial stress is equal to zero at the inner and outer surfaces. For both discs, it has positive value at the reference temperature ($T_0 = 0^\circ\text{C}$). However, with the increasing temperature it decreases gradually. The reduction

is also very much in Disc 2 when it is small for the other. Figure 7 illustrates tangential stress distributions along the radial direction of both discs. For both discs, the magnitudes of the tangential stresses are of lower values in the outer surface when they have higher values in the inner surface at the reference temperature, and after the middle section of the discs the stresses values come close to each other. But, with the increasing temperature the tangential stresses decrease at the inner surface whereas they increase at the outer surface for both discs. Figure 8 depicts radial displacement distributions along the radial section of both discs. The radial displacements increase less in the inner edge whereas increase more in the outer edge with the increasing temperature when they in the reference temperature adjacent to each other, approximately. If the both discs are compared with each other, it can be seen from the figure that the radial displacement values in Disc 2 are higher than those in Disc 1.

3.1 Radial Stress Plot

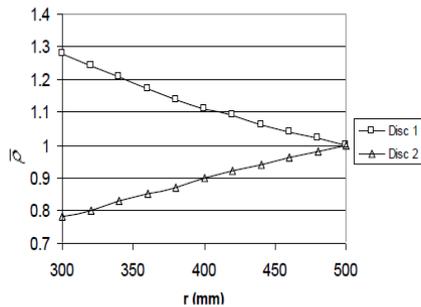
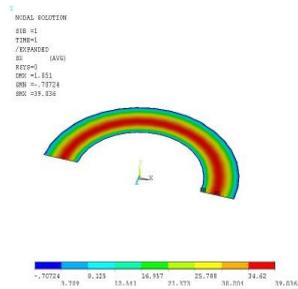


Figure 3. The normalized density distributions along the radial section of the discs.

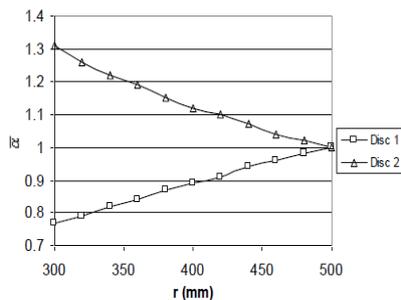


Figure 4. The normalized temperature expansion coefficient distributions along the radial section of the discs.

3.2 Tangential Stress Plot

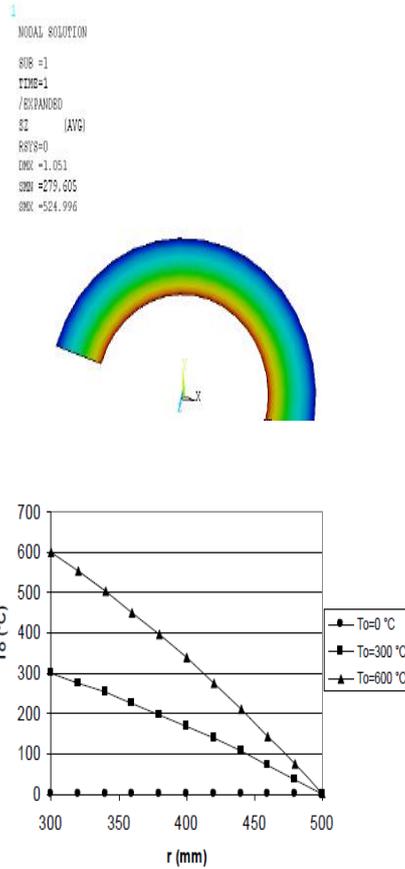


Figure 5. The temperature applied distributions along the radial section of the discs.

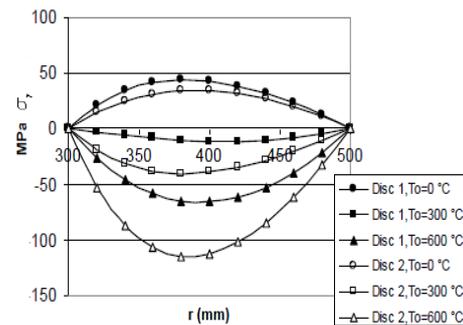


Figure 6. Radial stress distributions along the radial section of the discs.

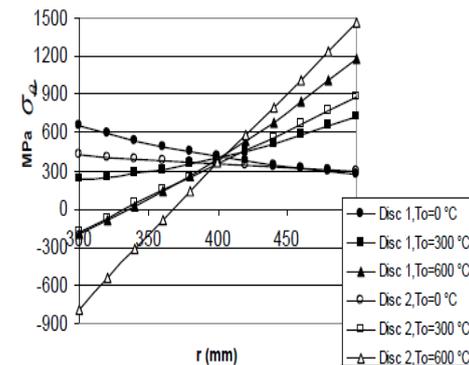


Figure 7. Tangential stress distributions along the radial section of the discs.

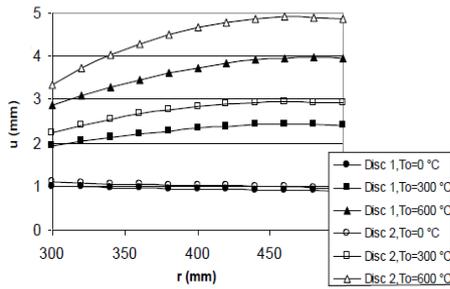
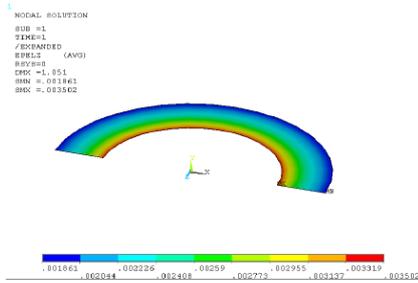
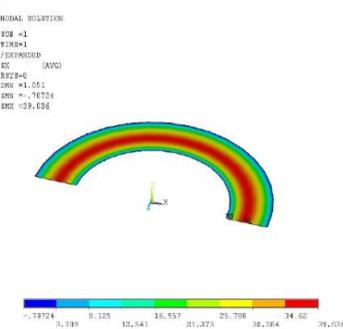


Figure 8. Radial displacement distributions along the radial section of the discs.

3.3 Tangential Strain Plot



3.4 Radial Displacement Plot



are found to be consistent with each other. stress analysis of the functionally graded discs:

- 1) Thermal expansion coefficient decreases when elasticity modulus and density increase, or this situation is diametrically opposite for functionally graded materials.
- 2) The tangential stress components are found to be highest at the inner surface but lowest at the outer surface for both Discs. They decrease at the inner surface whereas increase at the outer surface by increasing temperature.
- 3) The radial stress components decrease gradually along the radial section when the temperature is increased.
- 4) The magnitudes of the tangential stresses are higher than those of the radial stresses.
- 5) The analytical solution gives the radial displacement component at each point. The radial displacements increase more and more at the inner and outer surfaces by increasing temperature for both discs.

REFERENCE

- [1] Jahed H, Dubey RN. An axisymmetric method of elastic plastic analysis capable of predicting residual stresses. *J Press Vessel Technol* 1997;119:264–73.
- [2] Jahed H, Sherkatti S. Thermoplastic analysis of inhomogeneous rotating disk with variable thickness *Proceedings of the EMAS Conference of Fatigue, Cambridge, England April 2000.*
- [3] Vivio and Vullo Thermoplastic analysis of rotating at elevated temperatures. *Int J Press Vessels Piping* 2001;71(3):285–91.
- [4] Timoshenko SP, Goodier JN. *Theory of elasticity.* New York: McGraw-Hill; 1970.
- [5] Leopold WR. Centrifugal and thermal stresses in rotating disks. *ASME J Appl Mech* 1984;18:322–6.
- [6] Hosseini Kordkheili and Naghdabadi a semi-analytical thermoelasticity solution for hollow and solid rotating axisymmetric disks 1978;44:1–11

IV. CONCLUSION

The following conclusions are found from the thermal Table1. Comparison of analytical and numerical results in rotating isotropic homogeny disc.

procedure	surface	σ_{θ} (MPa)	U(mm)	ϵ_{θ}	$(\sigma_r)_{\max}$ (MPa)
Numeric	Inner	524.996	1.051	.00350	39.0360
	Outer	279.605	0.93062	.00186	
Analytic	Inner	525.252	1.0505	.00350	39.0371
	outer	279.188	0.93062	.00186	

Maximum radial stresses, tangential stresses, radial displacements and tangential strains values obtained from ANSYS commercial finite element analysis program and the present analytical study at the inner and outer surfaces of the only rotating isotropic homogeny disc are given in Table 1 and their numerical results are depicted in Figure 9. As seen in the table, the analytical results are compared with the numerical results which are obtained from ANSYS and they

Efficiency of General Insurance in Malaysia Using Stochastic Frontier Analysis (SFA)

Mohamad Arif Awang Nawi¹, Wan Muhamad Amir W Ahmad² and Nor Azlida Aleng³

(Department of Mathematics, Faculty of Science and Technology, Universiti Malaysia Terengganu (UMT)
21030 Kuala Terengganu, Terengganu, Malaysia)

ABSTRACT: General insurance comprises insurance of property against fire and burglary, floods, storms, earthquakes and so on. The purpose of the current study is to measure the relative efficiency of general insurance in Malaysia by using SFA for the year 2007 until 2009, consist of 26 general insurance companies by using the software FRONTIER to obtain the maximum likelihood (ML) and to get the relative efficiency. The finding showed that Oriental Capital Assurance Bhd (OCA) is at rank 1 for the three years. The 0.03975 value for the variance gamma (γ) parameter in this study is far from one, suggesting that all of the residual variations are not due to the inefficiency effects, but to random shocks. It can therefore, be concluded that the technical inefficiency effects associated with the production of the total profits by the input of the general insurance are very low.

Keywords: Efficiency, stochastic frontier analysis, specification of Battese and Coelli, general insurance.

I. INTRODUCTION

Insurance is a form of risk management in which the insured transfers the cost of potential loss to another entity in exchange for monetary compensation known as the premium. General insurance comprises insurance of property against fire and burglary, floods, storms, earthquakes and so on. It covers personal insurance as well as insurance against accidents and also covers health insurance and liability insurance which guards legal liabilities. Then again it covers other areas such as errors and omissions, insurance for professionals, credit insurance etc. Most general insurance policies are annual and the premium payment is in advance. No risk commences unless you have paid the premium. In some long term policies companies have the facility of collecting premiums periodically [2].

This study aims to measure the relative efficiency of 26 general insurance companies in Malaysia from 2007 to 2009. The main objectives identified are as follows:

- i. Measure the relative efficiency of performance for the general insurance industry.
- ii. Identify the most efficient general insurance companies based on relative efficiency scores.
- iii. Analyze the comparative efficiency of general insurance companies in Malaysia.
- iv. To estimate technical efficiency of general insurance by using Stochastic Frontier Analysis.

II. LITERATURE REVIEW

Fetcher et al. [10] has been using the Stochastic Frontier Analysis (SFA) to analyze the cost and efficiency of life insurance and general insurance in France using data from 1984 to 1989. From the study found that the average level of efficiency for life insurance is 30% and general insurance is 50%. Greene and Segal [11] have examined the relationship between cost inefficiency and profit for the life insurance industry in the United States. Profit is important for an insurance firm because of capital gains and determines the potential of a firm. Greene and Segal [11] distinguish the cost of efficiency using the stochastic frontier which this method using inefficient means for searching the diversity of firms and output. They also proposed that the cost efficiency in the life insurance industry will be strong when there is turnover and inefficiency occurs when the gain is measured by return on equity.

Fenn et al. [9] had use stochastic frontier analysis to estimate Flexible Fourier cost and profit functions for European insurance companies. They also adopt a maximum likelihood approach to estimation in which the variance of both one-sided and two-sided error terms is modeled jointly with the frontiers. This approach causes that simultaneously control for the impact of heteroscedasticity on the estimation of scale economies as well as estimating the effect of firm size and market structure on X-inefficiency. Separate frontiers are estimated for life, nonlife and composites companies and use data set of financial reports for the period 1995 to 2001. This provides technical and non-technical accounts at year-end for life, non-life and composite insurance businesses in 14 major European countries. The result showed that estimates most European insurers are operating under conditions of decreasing costs, and that company size and market share are significant factors determining X-inefficiency with respect to both costs and profits. Fenn et al. [9] also state that the larger firms and those with high market shares tend to have more cost inefficiency but less profit inefficiency.

Kasman and Turgutlu [19] investigates the technical efficiency of a sample of Turkish life insurance firms using the deterministic data envelopment analysis (DEA), chance-constrained data envelopment analysis (CCDEA) and stochastic frontier analysis (SFA) for the period 1999 to 2005. The main objective is to provide new information on the effect of methodological choice on the estimated efficiency by applying econometric and mathematical programming techniques to the same data set of Turkish life insurance firms. The empirical results show that the parametric and non-parametric methods provide similar rankings of firms but they differ significantly when the mean efficiency scores are

considered. From the results suggest that the stochastic structure of the CCDEA approach does not eliminate the fundamental differences between DEA and SFA. From the result, the three techniques suggest that there is a significant inefficiency problem in the Turkish life insurance industry over the sample period.

For the theoretical of stochastic frontier functions have not explicitly formulated a model for the inefficiency effects. Empirical papers, in which the issue of the explanation of the inefficiency effects has been raised include Kalirajan [14][15][16], Kalirajan and Flinn [17], and Kalirajan and Shand [18]. Their studies adopt a two-stage approach. The first stage involves the specification and estimation of the stochastic frontier production function and the prediction of either the inefficiency effects or the technical efficiencies of the firms involved. The second stage of the analysis involves the specification of a regression model for either the predicted inefficiency effects or the levels of technical efficiency of the firms in terms of various explanatory variables and an additive random error. The parameters of this second-stage inefficiency model have been generally estimated by using ordinary least-squares regression.

Study in Malaysia conducted by Shazali and Alias [23] review the performance of productivity and efficiency the life insurance industry for the community in Malaysia. This study is an attempt to measure productivity in the life insurance industry based on the method of Malmquist Non-parametric Index. The study found that although the insurance industry productivity is increase but the relative growth of life insurance is still low compared with the actual growth of the Malaysian economy. Just as the manufacturing sector, this sector's future growth depends on its ability to compete efficiently. The ability to provide efficient service is an important source of competitive advantage in the era of globalization. The study also found that the efficiency of technology development and contribute to the overall productivity in the industry.

III. SPECIFICATION OF STOCHASTIC FRONTIER MODEL

Berger et al. [6] and Berger and Humphrey [5] has introduced two techniques to measure efficiency. There are several econometric (parametric approach) and linear programming (nonparametric approach). The parametric approach has the advantage of allowing noise in the measurement of inefficiency. However, the approach needs to specify the functional form for production, cost or profit. Coelli [8] state that the non-parametric approach is simple and easy to calculate since it does not require the specification of the functional. The method for this study is Stochastic Frontier Analysis (SFA) by using the model of Battese and Coelli [3].

SFA is a way in economic modeling. Aigner, Lovell, and Schmidt [1], Meeusen and van den Broeck [21], and Battese and Cora [4] introduced the parametric approach to estimate stochastic production frontiers. These approaches specified a parametric production function and a two-component error term. One component, reflecting the influence of many unaccountable factors on production as well as measurement error, is considered "noise" and is usually assumed to be normal. The other component

describes inefficiency and is assumed to have a one-sided distribution, of which the conventional candidates include the half normal [1], truncated normal [24], exponential [21] and gamma [12][13] [24].

Battese and Coelli [3] assume a traditional random error (V_{it}) and a nonnegative error term (U_{it}) representing the technical inefficiency. Here, V_{it} is assumed to be independent and identically distributed, i.i.d $N(0, \sigma_v^2)$ and captures statistical noise, measurement error, and other random events (i.e., economic situations, quakes, weather, strikes and luck) that are beyond the company's control. The non-negative error term (U_{it}) captures the inefficiency and is assumed to be i.i.d as truncations at zero of the $N(\mu, \sigma_u^2)$. Also, V_{it} is assumed to be independent of the U_{it} . The model may be formed as follows:

$$Y_{it} = X_{it}\beta + (V_{it} - U_{it}) \quad i = 1, \dots, K; t = 1, \dots, T \quad (1)$$

Where Y_{it} is output of the i^{th} firm in the t^{th} time period; X_{it} is a $K \times 1$ vector of inputs of the i^{th} firm in the t^{th} time period; β is a $K \times 1$ vector of unknown parameters; V_{it} and U_{it} are assumed to have normal and half-normal distribution, respectively. This method can compile the efficiency of the insurance company according to its function and not using a specific distribution function. Features found in this method are suitable for measuring the efficiency of insurance companies because it will be arranged in the most efficient level. With the information obtained from this study can help rehabilitate the making of new policies for improvement and further enhance the growth of the general insurance industry in Malaysia. According to the equation, the model for a specific general insurance used in this study is:

$$\log Y_{it} = \beta_0 + \beta_1 \log x_{1it} + \beta_2 \log x_{2it} + \beta_3 \log x_{3it} + \beta_4 \log x_{4it} + \beta_5 \log x_{5it} + V_{it} - U_{it} \quad (2)$$

where,

Y_{it} = total profits of companies of the i^{th} company in the t^{th} time period

β = vector of unknown parameters to be estimate

x_{1it} = net investment income of the i^{th} company in the t^{th} time period

x_{2it} = management expenses of the i^{th} company in the t^{th} time period

x_{3it} = total liabilities and assets of the i^{th} company in the t^{th} time period

x_{4it} = annual premium of the i^{th} company in the t^{th} time period

x_{5it} = net claims paid by the company of the i^{th} company in the t^{th} time period

U_{it} = non-negative random variables, associated with

technical inefficiency of total profits of companies.

V_{it} = assumed to be independent and identically distributed (i.i.d) $N(0, \sigma_v^2)$ and captures statistical noise, measurement error, and other random.

Battese and Coelli [3] has proposed a stochastic frontier production function is defined for panel data on firms, in which the non-negative technical inefficiency effects are assumed to be a function of firm-specific variables and vary over time. The inefficiency effects are assumed to be independently distributed as truncations of normal distributions with constant variance, but means which are a linear function of observable firm-specific variables. The generalized likelihood-ratio test is considered for testing the null hypotheses, that the inefficiency effects are not stochastic or that they do not depend on the firm-specific variables.

The variance parameters are:

$$\sigma^2 = \sigma_v^2 + \sigma_u^2 \tag{3}$$

The maximum-likelihood method is applied for the estimation of the parameters of the model and the prediction of the technical efficiencies of the firms over time. This method gives more satisfactory results as more efficient than the method of *ordinary least squares (OLS)* [22]. Parameters (γ) must be in the range between 0 and 1. Parameters for the stochastic production function estimated using the maximum-likelihood estimation method and the calculation by using the Frontier Version 4.1c [7].

$$\gamma = \sigma_u^2 / \sigma^2 \tag{4}$$

So that $0 \leq \lambda \leq 1$.

Given the specification of the stochastic frontier model in Equation (1), the technical efficiency of production of the i^{th} firm given the level of inputs is defined as:

$$TE_i = \exp(-u_i) \tag{5}$$

So that $0 \leq TE_i \leq 1$ are inversely related to technical inefficiency [20]. In this study, the parameter γ is important because it facilitates the analysis of the efficiency of general insurance companies to be studied efficiently or not. The test statistic t and t distribution are used. Significance level used in this study are 0.05 and 0.01. The hypothesis of the study is as follows:

$H_0: \gamma = 0$ where, technical inefficiency of the insurance company investigated

$H_1: \gamma > 0$ where, technical efficiency of the insurance company investigated

IV. RESULTS ANALYSIS

This study discusses about the general insurance company where efficiency score indicates the highest value is the most efficient and reflects the company is able

to maximize the input which is very well without any problems. The total number of observations from this study is 26 firm operating in Malaysia general insurance industry over the period 2007 to 2009. This time period was chosen because the time had remained in a stable phase after a variety of the economic recovery process in Malaysia. The results showed an increase from year to year as shown in Table 1 for general insurance. This positive improvement show that the insurance industry has been in high demand among the people. It was found that all general insurance companies to perform as indicated when their efficiency is at 0.8 or 80% and above.

The result showed that *Oriental Capital Assurance Bhd (OCA)* is at rank 1 for the three years and it has a relative efficiency score higher than most other companies. Efficiency scores for OCA company starts from 0.94 in 2007 and increased to 0.97 in 2009. Score for the second position is BERJAYA and MUI CONTINENTAL. All two of these insurance companies get score as much as 0.92 in 2007, increased to 0.94 for year 2008 and 0.96 in 2009. Besides that, Pacific & Orient (P & O) insurance obtain a score of 0.91 in 2007 (rank 3). In 2008 the P & O insurance company showed an increase of 0.03 (0.94) and in 2009 the score was 0.96. The last position is PROGRESSIVE company obtained score 0.86 (rank 8) in 2007, 0.90 (rank 6) in 2008 and 0.93 (rank 5) in 2009. Mean of efficiency score of general insurance companies were increasing from year to year presented in Table 1. Every company indicates good performance by the year to manage the company's management expenses as well as possible in order to profit the whole companies.

Table 1: Relative Efficiency Score for General Insurance Companies from Year 2007 to 2009

	Insurance Companies	Efficiency Score			Rank		
		2007	2008	2009	2007	2008	2009
1	ACE	0.9	0.93	0.95	4	3	3
2	AGIC	0.9	0.93	0.95	4	3	3
3	AXA	0.89	0.92	0.95	5	4	3
4	BERJAYA	0.92	0.94	0.96	2	2	2
5	ETIQA	0.87	0.91	0.93	7	5	5
6	GREAT EASTERN	0.88	0.92	0.94	6	4	4
7	HONG LEONG	0.88	0.91	0.94	6	5	4
8	ING	0.89	0.92	0.94	5	4	4
9	JERNIH	0.88	0.91	0.94	6	5	4
10	KURNIA	0.9	0.93	0.95	4	3	3
11	LONPAC	0.89	0.92	0.94	5	4	4
12	MAA	0.88	0.91	0.94	6	5	4
13	MCIS ZURICH	0.9	0.93	0.95	4	3	3
14	MUI CONTINENTAL	0.92	0.94	0.96	2	2	2
15	MULTI-PURPOSE	0.91	0.94	0.95	3	2	3
16	OAC	0.9	0.93	0.95	4	3	3
17	OCA	0.94	0.96	0.97	1	1	1
18	P&O	0.91	0.94	0.96	3	2	2
19	PACIFIC	0.89	0.92	0.95	5	4	3
20	PROGRESSIVE	0.86	0.9	0.93	8	6	5
21	PRUDENTIAL	0.89	0.92	0.94	5	4	4
22	QBE	0.89	0.92	0.95	5	4	3
23	RHB	0.89	0.92	0.94	5	4	4
24	BH INSURANCE	0.88	0.91	0.94	6	5	4
25	TOKIOMARINE	0.9	0.93	0.95	4	3	3
26	UNILASIA GENERAL	0.89	0.92	0.94	5	4	4
Mean of efficiency score		0.89	0.92	0.95			

(Sources: Model Output Display by Battese and Coelli, 1992)

Figure 1 shows the relative efficiency of general insurance companies for a period of 3 years commencing from 2007 to 2009. From the diagram below, can conclude that the performance efficiency of the general insurance industry increased by an increase of 0.02 (2%) to 0.04 (4%) per year.

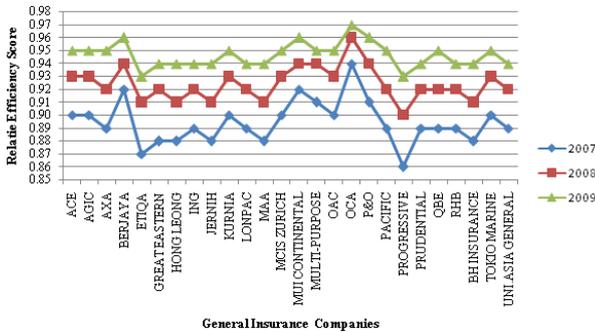


Fig. 1: Relative Efficiency Score for General Insurance Companies from Year 2007 to 2009

Table 2 presents the maximum likelihood estimate (MLE) for the parameters of the linear production function and related statistical tests results obtained from the stochastic frontier analysis. The estimated production function parameters indicated that net investment income, management expenses, total liabilities and assets, annual premium and net claims paid by the company. The estimated sigma-squared (σ^2) in this study is 0.02455 and it statistically different from zero at one percent. The result indicates that the one-sided error term dominates the symmetry error indicating a good fit and the correctness of the specified distributional assumptions.

The result obtained from testing the hypothesis of a general insurance company is calculated t statistic is 0.19281 and it is less than the critical value of t statistics 2.375 ($t = 0.19281 > t_{0.01} = 2.375$). Therefore, the null hypothesis is not rejected at significance level 0.01. This indicates is not significant relationship between net investment income, management expenses, annual premium and net claims paid by company. The 0.03975 value for the variance gamma (γ) parameter in this study is far from one, suggesting that all of the residual variations are not due to the inefficiency effects, but to random shocks. It can therefore, be concluded that the technical inefficiency effects associated with the production of the total profits by the input of the general insurance are very low. Nevertheless, the gamma which is statistically significant suggests that the traditional (OLS) function is not an adequate presentation.

Results of maximum likelihood estimation (MLE) found that the elasticity of the total profits for annual premium is the highest 0.9046. This means that with an increase 1% in input annual premium will increase by 0.9046% on the profitability of general insurance companies. For the input of the net investment income, the elasticity of the total profit is 0.2178. This means that 1% increase in net investment income input resulted in an increase of 0.2178% on the profitability of general insurance companies.

Table 2: Results of Maximum Likelihood Estimation for the General Insurance Industry

Parameters	Coefficient	t-ratio
Constant	$\beta_0 = 0.52997$	0.29619
Net investment income	$\beta_1 = 0.21778$	3.6814**
Management expenses	$\beta_2 = -0.03655$	-0.6083
Total liabilities and assets	$\beta_3 = -0.08334$	-1.0769
Annual premium	$\beta_4 = 0.90460$	16.7974**
Net claims paid by the company	$\beta_5 = 0.02082$	0.59358
Sigma-squared		
$(\sigma^2 = \sigma_{\mu^2} + \sigma_{\epsilon^2})$	$\sigma^2 = 0.02455$	5.95461**
Gamma		
$(\gamma = \sigma_{\mu^2} / \sigma^2)$	$\gamma = 0.03975$	0.19281
Log Likelihood Function	35.65203	

Significance level, 0.01**, 0.05*

V. CONCLUSION AND DISCUSSION

This study focuses on stochastic frontier analysis approach (SFA) that involves econometric methods used to analyze the efficiency of general insurance companies in Malaysia. Battese and Coelli model [3] is used to obtain the relative efficiency of general insurance companies from year 2007 to 2009. From the study, the relative efficiency for general insurance companies have been increasing from year to year. Companies who posted scores the highest relative efficiency for general insurance is *Oriental Capital Assurance Bhd (OCA)* is at rank 1 for the three years and it has a relative efficiency score higher than most other companies. Efficiency scores for OCA company starts from 0.94 in 2007 and increased to 0.97 in 2009. In addition, the efficiency performance of the general insurance industry increased by an increase of 0.02 (2%) to 0.04 (4%) by year. The 0.03975 value for the variance gamma (γ) parameter in this study is far from one, suggesting that all of the residual variations are not due to the inefficiency effects, but to random shocks. MLE found that the elasticity of the total profits for annual premium is the highest 0.9046 and the second highest is net investment income where the elasticity of the total profit is 0.2178. The results of this study also assessed by looking at the efficiency score. According to the rank of efficiency included in this study can help people in selecting and evaluating general insurance companies that have good performance. This study also will help the management and administration of insurance firms involved in making and improves the weaknesses, such as formulating business strategy or marketing strategy to attract customers who can benefit the firm. This study can also be used as a benchmark in determining the efficiency of insurance companies in Malaysia according to the appropriate model.

REFERENCES

[1] Aigner, D. Lovell, C. A. K & Schmidt P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometric* 6: 21-37.

[2] ArticleWorld. May 2012. General Insurance. Available at http://www.articleworld.org/index.php/General_insurance.

[3] Battese, G. E., & Coelli, T. J. 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis* 3: 153-169.

- [4] Battese, G. E., Corra, G. 1977, Estimation of a production frontier model: with application to the pastoral zone of eastern Australia, *Australian Journal of Agricultural Economics* 21, 167-179.
- [5] Berger, A.N. & Humphrey, D.B. (1997). Efficiency of financial institutions: international survey and directions for future research. *European Journal of Operational Research* 98: 175–212.
- [6] Berger, A.N., Hunter, W.C. & Timme, S.G. (1993). “The Efficiency of Financial Institution: a Review and Preview of Research Past, Present and Future”, *Journal of Banking and Finance*, vol. 17, no. 2-3, 221-250.
- [7] Coelli, T.J. 1996. A computer program for stochastic frontier production and cost function estimation. frontier version 4.1: CEPA working paper 96/97, Department of Econometrics, University of New England.
- [8] Coelli, T.J. 2004. “Efficiency and Productivity measurement: an overview of concepts, terminology and methods”, Paper presented at the short course on “Productivity and Efficiency. Measurement Methods with applications to infrastructure industries”, organised by University of Queensland, Brisbane, Australia, 25-27.
- [9] Fenn, P., Vencappa, D., Diacon, S., Klumpes, P. & O’Brien, C. 2007. Market structure and the efficiency of European insurance companies: A stochastic frontier analysis. *Journal of Banking and Finance* 32: 86-100.
- [10] Fetcher, F., Perelman, S. Kessler, D. & Pestieau P. 1993. Productive performance of the French insurance industry. *Journal of Productivity Analysis* 4: 77-93.
- [11] Greene, H. W. & Segal, D. 2004. Profitability and efficiency in the U.S life insurance industry. *Journal of Productivity Analysis* 21: 229-247.
- [12] Greene, W.H., 1980a, Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics* 13, 27-56.
- [13] Greene, W.H., 1980b, On the estimation of a flexible frontier production model. *Journal of Econometrics* 13, 101-115.
- [14] Kalirajan, K.P. (1981), "An Econometric Analysis of Yield Variability in Paddy Production", *Canadian Journal of Agricultural Economics* 29, 283-294.
- [15] Kalirajan, K.P. (1982), "On Measuring Yield Potential of the High Yielding Varieties Technology at Farm Level", *Journal of Agricultural Economics* 33, 227-236.
- [16] Kalirajan, K.P. (1989), "On Measuring the Contribution of Human Capital to Agricultural Production", *Indian Economic Review* 24, 247-261.
- [17] Kalirajan, K.P. and J.C. Flinn (1983), "The Measurement of Farm-specific Technical Efficiency", *Pakistan Journal of Applied Economics* 2, 167-180.
- [18] Kalirajan, K.P. and R.T. Shand. 1989. “A generalised measure of technical efficiency”. *Applied Economics*, 21: 25–34.
- [19] Kasman, A. & Turgutlu, E. 2007. A comparison of chance-constrained data envelopment analysis and stochastic frontier analysis: An application to the Turkish life insurance industry.
- [20] Khairo, S.A. and G.E. Battese. 2005. A study of technical inefficiencies of maize farmers within and outside the new agricultural extension program in the Harari region of Ethiopia. *South African Journal of Agricultural Extension*. Trangie Agricultural Research Centre. University of New England. 34 (1): 135-150.
- [21] Meeusen, W., van den Broeck, J., (1977). Efficiency estimation from Cobb-Douglas production function with composed error. *International Economic Review* 18, 435-444.
- [22] Richmond, J. 1974. Estimating the efficiency of production. *International Economic Review* 15: 515-521.
- [23] Shazali, A.M Alias, R. 2000. Productivity and efficiency performance of the Malaysian life insurance industry. *Jurnal Ekonomi Malaysia* 34: 93-105.
- [24] Stevenson, R.E., 1980, Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics* 13, 57-66.

Designing BT Antenna for RFID Systems

Ahmad Bayat

Department of Electrical Engineering , K.N.Toosi University of technology, Tehran, Iran

Abstract: This paper is about designing and optimizing a new antenna employed for RFID applications at microwave frequencies. The structure is designed in a way that it resonates at 2.45GHz, which is one of the RFID frequencies. The new patch is derived from BT(Bow-Tie) antenna. It results in a reduction of the size by more than 70%, as compared to monopole patch, alongside with having an acceptable gain and directivity. Based on a model of triangular antenna, a correspondent resonant model is presented. Simulation results achieved using our method show that the antenna and its model have the same resonant frequency at 2.45GHz. However, there was a little difference in Bandwidth.

Key Terms: Radio Frequency Identification RFID, Tag, Readable Range R, BT Antenna

I. INTRODUCTION

In recent years, microstrip antennas are being used for fine tuning light weight, low cost manufacturing, reliability and simplicity of the structure, making them widely applicable in research and engineering applications have been used.

Bow tie microstrip antennas are designed for wireless LANs operating frequency at 2.45 GHz applications. Visual Patch combines fantasy bow in two triangular patches, which are built on a single substrate. Figure (1) shows a ribbon bow that is a bow tie microstrip antenna. Bow tie microstrip antennas are attractive options because of their compact nature than rectangular patch for communication. Growing demand for compact wireless communication equipment increases the need for further research to more antenna option and this demand has led many researchers to the field of microstrip antennas bow tie option. However, few previous studies have addressed the analysis of this type of antennas and empirical formulas have been proposed for the frequency cohort of this new Geodesy.

To reach integrated transmit / receive modules, microstrip antennas should be mounted on a board that is consistent MMIC structures. Moreover, changing a strip line on a board to a coaxial cable or antenna feed waveguide is difficult. However, flat panel antenna can solve these problems, but they have limited bandwidth. Up to now broadband microstrip antennas have been studied. Recently printed dipole antennas and bow tie antennas are presented. Regardless of the good results of the bandwidth and radiation pattern of the antenna, microstrip structures are examined.

Microstrip lines widely used in manufacturing transistor microwave amplifiers because it is not difficult to make and it can also be used as inductors and capacitors.

It should be noted that the lines usually are used at high frequencies because its length gets long in the low frequency and making circuit will be impractical [13].

Microstrip line is a line that has a conductive strip of land and a home which are separated by a dielectric medium.

Figure (1) shows the structure of a Microstrip. Dielectric material acts as a medium between the conductors and the earth. The typical dielectric substrate, including two contact points with alumina and silicon, makes the dielectric wavelength higher and thereby, the whole circuit will be smaller.

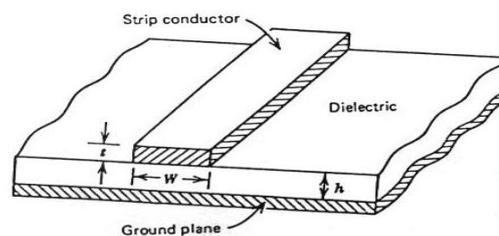


Figure 1. the physical structure of a Microstrip line

All lines of magnetic field in microstrip are not in bed. Therefore it cannot be considered a pure TEM because it has such a structure which is not true in the TEM.

Therefore, when choosing the antenna, some points should be considered. First, different types of antenna should be used in the reader and the tag for different applications. Second, the antenna should be small, easy to make and also cheap. Third, the RF characteristics of the antenna should be acceptable regarding impedance, gain and directivity.

On one hand, reducing the size of the antenna results in a small tag which can be placed anywhere, without occupying a large space. On the other hand, one of the main features of RFID systems is the readable range R, which is the maximum distance that the Interrogator can read from the tag. This distance basically depends on the gain of both the Tag antenna and the reader antenna, as represented in equation 1 [2]:

$$R = \frac{\lambda_0}{4\pi} \sqrt{\frac{P_1 \cdot G_{reader}^2 \cdot G_T^2}{P_3}} \quad (1)$$

With: P_3 = Power received in the tag

P_1 = Power transmitted by the interrogator

G_{reader} = The Gain of interrogator antenna

G_T = The Gain of Tag antenna

A new small tag antenna is proposed in this article. The patch is designed based on a bow tie antenna, and is simulated using Advanced Design System (ADS 2009). The significance of this new patch is that it results in the reduction of the size of the original antenna that resonates at 2.45GHz by more than 70%. I simulated an electrical model

of the structure to validate our work. Then I compared it with the ADS results of the physical patch.

II. A BT SMALL ANTENNA

A. Geometry of triangular antenna

The geometry of an equilateral triangular microstrip patch on a dielectric substrate with a ground plane is presented in Figure (2). The antenna is mounted on a substrate material with the thickness of $h=3.2\text{mm}$, a dielectric constant of $\epsilon_r = 2.6$ and the loss tangent of $(\tan\delta =) 0.002$.

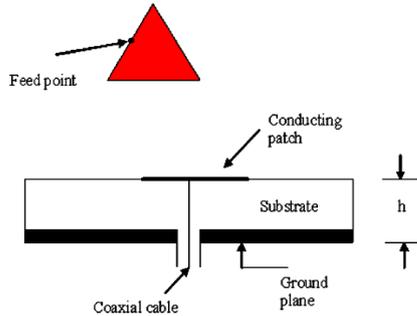


Figure 2. Triangular Antenna structure

B. Electrical model of triangular patch

The recent studies on microstrip patch have shown that the triangular patch has a radiation characteristic similar to a rectangular patch, but with reduced dimensions [3], [5], [6]. Due to this reason and in order to study our antenna, the triangular electrical model can be replaced by its equivalent electrical rectangular model, as shown in figure (3).

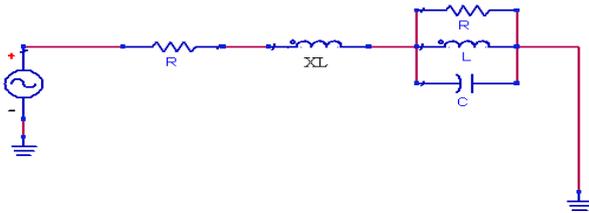


Figure 3. Electrical model of Triangular antenna

The formula presented in [4] is applied to every triangle in order to calculate the parameters of the model.

$$Z_{in} = R + jX \quad (2)$$

$$Z_{in} = \frac{R}{1 + Q_T^2 \begin{bmatrix} f & -f_r \\ f_r & f \end{bmatrix}} + j \left[X_L - \frac{RQ_T \begin{bmatrix} f & -f_r \\ f_r & f \end{bmatrix}}{1 + Q_T^2 \begin{bmatrix} f & -f_r \\ f_r & f \end{bmatrix}^2} \right]$$

Resonant resistance R :

Employing equation (3), the resonant resistance is calculated, as in [4]:

$$R = \frac{Q_T H}{\pi f_r \epsilon_{dyn} \epsilon_0 A} \cos^2 \left(\frac{\pi x_0}{a} \right) \quad (3)$$

f_r : Resonant frequency

Q_T : Quality factor

ϵ_{dyn} : Dynamic permittivity

x_0 : the distance of the feed point from the edge of the patch.

a : length of triangle

H : thickness of dielectric

A : air of triangle

$$Q_T = \left(\frac{1}{Q_R} + \frac{1}{Q_C} + \frac{1}{Q_D} \right)^{-1} \quad (4)$$

$$Q_R = \frac{c_0 \sqrt{\epsilon_{dyn}}}{4 f_r H} \quad (5)$$

$$Q_D = \frac{1}{\tan(\delta)} \quad (6)$$

$$Q_C = \frac{0.786 \sqrt{f_r Z_{a0}} (W) H}{P_a} \quad (7)$$

Q_R : Radiation quality factor

Q_D : Losses in the dielectric

Q_C : Losses in the conductor

$$Z_a(W) = \frac{60\pi}{\sqrt{\epsilon_r}} \left[\frac{W}{2H} + 0.441 + 0.082 \left(\frac{\epsilon_r - 1}{\epsilon_r^2} \right) + \frac{(\epsilon_r + 1)}{2\pi\epsilon_r} \left(1.451 + \ln \left(\frac{W}{2H} + 0.94 \right) \right) \right] \quad (8)$$

The impedance of an air filled microstrip line is as below.

$$Z_{a0}(W) = Z_a(W, \epsilon_r = 1)$$

$$P_a = \frac{2\pi \left(\frac{W}{H} + \frac{W/(\pi H)}{W/2H + 0.94} \right) \left(1 + \frac{H}{W} \right)}{\left(\frac{W}{H} + \frac{2}{\pi} \ln \left(2\pi \exp \left(\frac{W}{2H} + 0.94 \right) \right) \right)^2} \quad (9)$$

$$\frac{W}{H} > 2$$

$$\epsilon_{dyn} = \frac{C_{dyn}(\epsilon)}{C_{dyn}(\epsilon_0)} \quad (10)$$

$$C_{dyn}(\epsilon) = \frac{\epsilon_0 \epsilon_r A}{H \gamma_n \gamma_m} + \frac{1}{2\gamma_n} \left(\frac{\epsilon_{reff}(\epsilon_r, H, W)}{c_0 Z(\epsilon_r = 1, H, W)} - \frac{\epsilon_0 \epsilon_r A}{H} \right) \quad (11)$$

$$\gamma_j = \begin{cases} 1, & j = 0 \\ 2, & j \neq 0 \end{cases}$$

$$Z(W, H, \epsilon_r = 1) = \frac{377}{2\pi} \ln \left(\frac{f(W/H)}{w_h} + \sqrt{1 + \left(\frac{2}{W_H} \right)^2} \right) \quad (12)$$

$$f(W/H) = 6 + (2\pi - 6) \exp \left(- \left(\frac{30.666}{W_H} \right)^{0.758} \right) \quad (13)$$

Capacitance C, Inductance L:

For calculating the capacitance C, the formula of C_{dyn} is used.

To calculate the inductance L, we have:

$$\omega_{res} = 2\pi f_r \tag{14}$$

$$\omega_{res} = \frac{1}{\sqrt{LC}} \Rightarrow L = \frac{1}{\omega_{res}^2 C} \tag{15}$$

Inductive reactance of Coax:

Inductive reactance of Coax is determined using the following equation:

$$X_L = \frac{377 fH}{c_0} Ln\left(\frac{c_0}{\pi f d_0 \sqrt{\epsilon_0}}\right) \tag{16}$$

The model was simulated, then compared to the physical patch. Its results are shown in figure (4). The triangular patch and its electrical model show a resonant frequency of about 2.45GHz. However, they have different bandwidths, which is because of the difference between calculated and simulated losses. The gain and directivity are 8.1dB and 8.2dB, respectively.

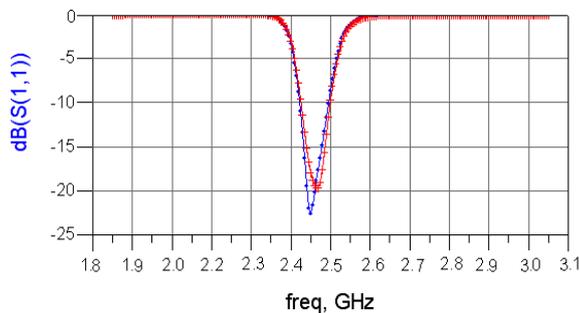


Figure 4. Return loss of both physical and model of triangular patch

C. Geometry of proposed bow tie antenna

The proposed patch consists of four triangle patches disposed on planar structure as shown in figure (5).

This structure is inspired by the bow tie patch antenna, in which only one transmission line links the four triangles, and the patch is mounted on a dielectric with a thickness of $h = 0.65\text{mm}$ and the permittivity of $\epsilon_r = 2.3$.

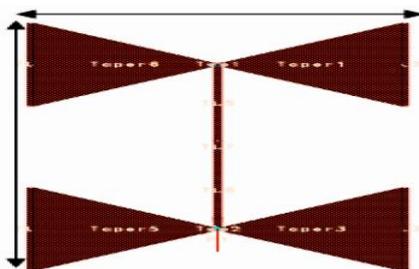


Figure 5. Figure 5. Proposed bow tie antenna structure

Figure 6.

D. Electrical model

An electrical model was developed for analyzing this antenna, which was inspired from the electrical model of the triangular patch described by Nasimuddin and A. K. Verma in [3]. In fact, they substituted the triangular patch by its equivalent rectangular patch and built the triangular electrical model. The transmission line was modeled by an RLC circuit. Figure (6) represents this electrical model.

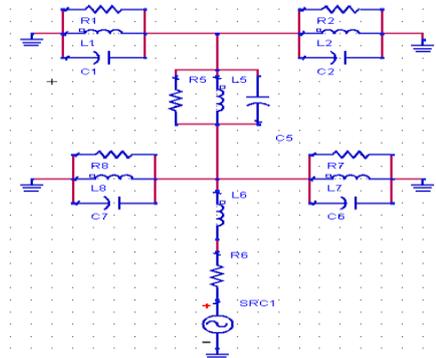


Figure 7. Electrical model of proposed bow tie antenna The RLC parameters were determined by means of the formulas developed in section 1.

III. RESULTS AND DISCUSSION

E. Dimensions

The patch proposed in this paper is designed in a way that it resonates at 2.45GHz (one of the RFID frequencies). The typical antenna is of the dimension of about $5.8\text{cm} \times 5.8\text{cm}$. The dimension of the antenna presented here is about $10.4\text{mm} \times 6.2\text{mm}$, which is more than 70% smaller than the typical antenna. It leads to a very small tag with a good quality.

Return loss

Figure (7) shows the return loss with the resonant frequency of 2.45 GHz. The antenna represents a band width of about 80 MHz which is satisfactory for RFID system. Also, the magnitude $|S_{11}|$ attained 25dB.

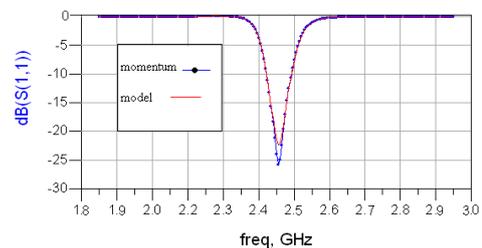


Figure 8. Return loss of both physical and model of proposed bow tie antenna

F. Gain and Directivity

Since the readable range, R, depends on Gain and Directivity, it should be assured that suitable ones are gained. In the bow tie antenna proposed here, the gain and he directivity are 2.1dB and 5.2dB respectively which are acceptable for an RFID system.

Finally, by comparing the physical patch and electrical model, it can be seen that the resonant frequency is the same. Furthermore, there is a little difference in band width which is due to the difference between losses in theory and real.

IV. CONCLUSION

Reducing the size of the antenna and the tag is greatly favored for many applications. Hence, many techniques are applied for reaching to a very small patch. Reducing the antenna size by more than 70% in this study is a very important achievement. Ameliorating antenna parameters and decreasing antenna cost are the priorities in our future research.

REFERENCES

- [1] Amir Galehdar, David V. Thiel, and Steven G. O'Keefe "Antenna Efficiency Calculations for Electrically Small, RFID Antennas", IEEE Antenna Wireless Propagation Letters, Vol 6, pp 156-159, 2007
- [2] Klaus Finkenzeller, "RFID HANDBOOK," Second Edition, 2003
- [3] Nasimuddin and A. K. Verma Amir, "Fast and accurate model for analysis of equilateral triangular patch antenna," Journal of Microwaves and Optoelectronics, Vol 3, pp 99-110, April 2004
- [4] Celal YILDIZ, Kerim GÜNEY, "Simple model for the impedance of rectangular microstrip antenna", journal of engineering sciences, pp 733-738, 1998
- [5] J.George, M.Deopukumar, C.K.Aanadan, P.Mohanan and K.G.Nair, New compact microstrip antenna, Electron lett vol.32, No.6, PP. 508-509, Mar.1996.
- [6] [6] Xu-Pu zhang, shum-Shi zheng, Resonant frequency and dual-frequency operation of a bow-tiemicrostrip antenna, IEEE AP-S Int Symp, vol.2, PP.60-63, June 2002.

Automation Engineering Design of Structures and Facilities

Shodmonkul Nazirov, Sabir Yakubov

Division of Informatics, Tashkent University of Information Technologies, Uzbekistan

Abstract: *The basic features of design optimization problems of engineering structures and buildings. The requirements for mathematical models optimization. It is shown that the process of optimal structures and facilities designing is a process of feedback control, which can be represented as a hierarchy of decision making. The main organization and functioning of a software program and CAD principles are developed.*

Keywords: *optimization, system analysis, design automation, mathematical model, mathematical and software, mathematical programming.*

I. Introduction

In the formulation of optimization problems in the engineering structures and buildings design an ambiguous interpretation (understanding) of Systems Analysis (approach) can be found: on one hand - this is really an analysis of any existing system, on the other - the formation of the system parameters to achieve the goals. In reality, the two positions go hand-in-hand, because you cannot create a system ,that provides goals, without analyzing the contents and determining the actual processes , that lead to the desired result. Systems analysis provides the conditions for joint optimization both structural parts of the system (subsystems) and the system as a whole, as well as computer software. The ultimate goal of systems analysis in the design is an actual design of the system, its subsystems and components for optimum efficiency and economy. Despite the fact that there are no well-defined rules in systematic analysis , the basic features are adequately disclosed in [1].

II. Main design processes of engineering structures and facilities

Taking into account the specifics of the design process of engineering design and construction tasks , the main features of our approach can be displayed in the following :-

I. As an optimized projected engineering design or construction, we take determined appropriate performance of the functions, the complex elements, endowed with specified properties and having links to the abstract and the external environment systems [I-III].

In this complex in the research process we can attach to each element the desired properties without regard to actual performance, to identify the possible contribution of these properties in the studied process and, therefore, justify the requirements for the prospective deal with this item. In practical optimization problems it is assumed that the properties of elements and their functional and technical specifications are known and therefore the functioning of the processes is considered in the field of permissible (taking into account the accepted limit) solutions of the systems. Both in first and second, as well as in the case of the software (development of algorithmic systems) evaluation of the complex is considered with taking into account the totality of known processes and phenomena and the relationship between them. All this highlights such features of the model of designed engineering structures and buildings that contribute to the elucidation of the functioning of this complex in order to select the least weight or cost.

The most significant is that in all cases, the system includes the concept of a whole consisting of interrelated, interacting and interdependent parts. The properties of these parts depend on the system as a whole, and the properties of the system – on the properties of its parts.

II. The place for specific design and engineering projected construction should be determined in the overall structure of the other systems. The systems approach requires a reasonable allocation of the designed system in general systems structure to maintain the normalizing parameters, dividing it into subsystems.

The design or construction is considered as an independent object of study and optimization, but with the necessary information exchange with the adjacent and external systems, and within it, between the subsystems.

The chosen general structure of systems should clearly delineate the boundaries of the system and contribute to selection (structuring) of its subsystems, which in size are available for research and homogeneous in the description. This provides links to the organization on each successive level, the descent of the system to the individual elements from the top - down with the subsequent transfer of the resulting aggregated data to the top (bottom-up).

Both the overall structure of compensation systems and subsystems of structures and facilities must have inherent properties of integrity: the changes that have arisen in any of their parts, impact on other parts, and on their entirety.

III. Engineering design or construction are presented in the form of the model. Complex systems designing, that the engineering designs and constructions are, requires knowledge of the quantitative and qualitative patterns of the system and its individual elements behavior , depending on the nature of the changes of multiple factors (parameters).

The model should be similar to the original, but differ from it. Its distinctive features are manifested in the fact that it is subjected to such transformations in the right direction, which is not possible in the direct study of the original.

Ultimately, the choice of method is determined by taking into account many considerations, the latter role of which belongs to the convenience of the treatment algorithm, the duration of the account, etc. It is also clear that the tasks require

informal activities, opportunities to intervene in the process of counting and to obtain interim results for the implementation of a dialog mode.

IV. To assess the quality of designed structures and buildings solutions a set of indicators is selected. As a rule, the purpose of system analysis is, that for all possible characteristics of external relations, to achieve the best possible (optimal) solutions of designed structures and buildings according to their structural, economic and other indicators. However, the optimum and optimality is not an absolute concept, they require a precise definition of optimality criteria, i.e. the main features on which the effectiveness of different solutions are compared. Solution, the best in one setting and one criterion, may be far worse in other circumstances and other criteria. Optimizing on one criterion (sub optimization), mostly for technical systems is aimed at reducing the cost (in this study the objective function is the weight of the structure).

V. The analysis of the model structures and buildings should be transferred to the real systems. To transfer the solutions to the real object requires confidence in the adequacy of the solution. The adequacy is estimated by the analogy of a real object and the model properties according their main characteristics. Adequacy is achieved if the model fully reflects the stress-deformed state (VAT) of actual projected structures and facilities.

Listed and taken to the execution, the main provisions of the systems approach are characterized by a basic framework of the method, but the efficiency of its use depends entirely on the way of their implementation. In order to systematize and generalize the information about the main symptoms of systemic analysis that contribute to the representation of disparate data in an orderly manner with a smaller number of significant variables, you must:

- systemize relationship between the systems, designed to maintain the normalized parameters;
- analyze multiple baselines, to find a form of generalization, suitable for determining the VAT system conditions classification;
- identify the classification structure or structures, contributing to a focused selection of competing options;
- determine the principles of decomposition of systems, based on the analysis of their aggregate as a whole;
- formulate a basic framework for constructing a mathematical model of the design or construction;
- classify optimization problems, arising in practice, research and design.

The purpose of the design process is to develop technical documentation, required for the production of design objects and based on both priori (the original) and posteriori (additional) information, obtained in the design. Thus, the design is the process of creating a prototype of an object, indispensable for the object manufacture. Design is essentially a process of feedback control (Fig. 1). Terms of Reference (TOR) produce inputs or outputs, which are compared with the results of the design, and if they do not match, the design cycle repeats itself as long as the error (deviation from the set of technical requirements) is not within limits.

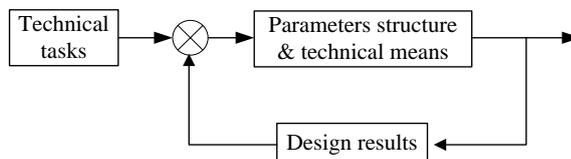


Fig. 1. The scheme design

The design process is system design, i.e. a set of interacting with each other designers and means, required for the hardware. In essence, the system design can be considered as a complex man-machine multi-loop, multi-dimensional control systems with feedback, requiring the collection, transmission, processing and using information to achieve the design objectives. They should be subject to one or another criterion of optimization, for example, the criterion of minimum length or maximum performance on a limited budget, or the fastest designed system payback criterion, etc. To reduce design time it is necessary to increase the speed of transmission of useful signals, and to prevent interference, i.e. signals, not carrying useful information. The incorrect or inaccurate intermediate results, or an unfortunate choice of the structure of the system design, when the signals, necessary for making decisions on any low level, get to the upper levels, where they can be not only useless but harmful, may give such interference. On this basis, we can conclude that a systematic approach, principles and methods of control theory are of great interest for the rational organization of the design.

"Computer-Aided Design" (CAD) was used in all cases where a computer was used for calculations, related to the design. Currently Apr for engineering structures and buildings is associated with the new stage of development - the creation of computer-aided design (CAD), primarily designed to meet the challenges of scientific research and conceptual, partly technical, design. CAD can be considered as a system of feedback control. Therefore, we can give the following definition, emphasizing the "management" aspect of CAD. CAD is a man-machine or an automated process control system design of engineering structures and buildings and technical documentation, required for the manufacture of the designed object. Common to all CAD systems is that they are, as already was indicated above, regardless of the design object, may themselves be regarded as the automated process control system for the production of technical documentation. Therefore, the development of a general theory of automated control systems is at the same time the development of theoretical foundations of any CAD objects.

III. Steps to create automated design of optimal structures and facilities

Thus, the essence of computer-aided design of optimal engineering structures and buildings comprises:

1. The development of the optimized mathematical model of the object (the projected design and construction) and the external environment on the a priori available information and the results of the identification of existing facilities or physical models.

2. Structural synthesis on the basis of analysis of the design projected features and construction, determination of the necessary information flows, taking into account the possibility of decomposition of the system and distribution management functions to appropriate levels.

3. Structural synthesis of algorithmic (automated) optimization of engineering structures and buildings system.

4. Definition of the computer interaction modes.

5. Analysis of the results of computer simulation.

6. Clarifying and making amendments to the TOR.

7. Documentation.

Let's call computer-aided design of engineering structures and buildings, carrying out the first five of the above points, a system of automated synthesis.

Certainly the development of CAD engineering structures and buildings must be guided by standard operating system.

One of the most important bases to create CAD software application for optimization of engineering structures and buildings is the method of computational mathematics. However, the "adaptation" of these methods to a form that is possible and convenient for implementation of the computing means, is a complex and laborious process that represents the life cycle of a software product. The most important step in this process is the development of algorithmic software [2].

Once you made up the algorithm of solutions of a problem, the programming process, i.e. coding algorithm in terms of selected high-level language, or directly in terms of machine instructions starts.

Created (designed) software package (RFP) for the optimization of engineering structures and buildings should have the following features:

- be built in a modular (sub systems) principle;
- have some flexibility in the ratio of hardware and software operating environment;
- availability for poorly prepared in the field of computer science professionals;
- have friendly means of advanced problem-oriented dialogue;
- allow interface with tool support systems: databases, graphics systems, databases;
- allow modification and expansion.

IV. The scheme for obtaining the optimal solution

Let's consider the integrated scheme for optimal solutions. To create a calculation model, taking full account of all the properties of a real object (the deformed system) is fundamentally impossible. The art of choosing the computational model is to identify the basic properties of a real object. After selecting a design model, a mathematical model for describing the strain and stress states, the dynamic processes, etc., connecting the incoming parameters should be composed. Availability and preliminary analysis of a mathematical model, describing the state of a deformable system, allows giving the task of the problem optimization and finding an effective mathematical tool optimization. To estimate the received solution (project) it is necessary to define an optimality criterion (the criterion of excellence, quality criteria, the criterion of efficiency, etc.). The objective function $C(x)$ is a mathematical account of the criterion of optimality. Using of the criterion of optimality, limitations of the mathematical model describing the state, the mathematical apparatus of optimization (optimization techniques) allows you to find the optimal solution, which describes the projected structure (object model).

V. A mathematical model of optimization of cylindrical shells by weight

The problem of optimizing the design put a mathematical programming problem: it is necessary to determine the vector X (x_1, x_2, \dots, x_n) of x_i ($i = 1, \bar{n}$) optimized parameters, giving the objective function $F(x)$ extreme (for definiteness, we take min), subject to restrictions on the parameters $a_i \leq x_i \leq b_i$, ($i = 1, \bar{n}$); and functional limitations $f_j(x) \leq 0$ ($j = 1, \bar{m}$) [3]. This problem can be written as

$$\left. \begin{array}{l} F(X) \rightarrow \min \\ f_j(X) \leq 0 \quad (j = 1, \bar{m}) \\ a_i \leq x \leq b_i \quad (i = 1, \bar{n}) \end{array} \right\} \quad (1)$$

For example, let's consider the problem of optimization of cylindrical shells by weight. We assume the objective function as

$$F(x) = \int_{\alpha} \int_{\beta} h(\alpha, \beta) R d\alpha d\beta \quad (2)$$

For open shell type codes are given:

- a) the boundary conditions;
- b) the length of the overlap- a;
- c) the width of the overlap;
- d) the material of the shell;

E modulus of allowable elasticity; σ - allowable stress; γ - gravity; ν - Poisson's ratio; [U] - allowable transfer (if required limitations on the stress; implementation of strength and stiffness);

- e) - a system of external loads;

e) - other restrictions (such as design, technological, etc., if you want to meet them.).

Optimize able parameters are parameters that determine the variation of shell thickness, angle of the shell, which determines the degree of its steepness.

VI. The results of computer simulation

At present we have developed the engineering constructions optimization software on the base of global search [3]. It is realized in visual program system Delphi 7.

Task 1. Optimizing of cylinder shell ,of rectangular shape, the whole shape hinge supported, under uniform average load of q intensity.

The shell thickness is constant $h = const$.

The physical parameters of shell material: $E = 2 \cdot 10^6 \text{ kg/cm}^2$; $[\sigma] = 2000 \text{ kg/cm}^2$; $\nu = 0,5$ The shell geometrical parameters:

$a=150\text{sm}$
 $b=100\text{sm}$

Load $q=1 \text{ kg/sm}^2$.

Optimized parameters h, β_0 .

Parameters restrictions :

$$\frac{\pi}{10} \leq \beta_0 \leq \pi;$$

$$0,1\text{cm} \leq h \leq 3\text{cm}$$

Minimized function is the cross-section area

$$S = R \cdot h \cdot \beta_0$$

Construction restrictions:

$$\sigma_i \leq [\sigma]$$

Where σ_1 -stress intensity is calculated according to:

$$\sigma_i = \sqrt{(\sigma_{11} - \sigma_{12})^2 + (\sigma_{22} - \sigma_{33})^2 + (\sigma_{33} - \sigma_{11})^2 + 6\tau_{23}^2} \quad (3)$$

Stress values $\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{23}, \tau_{23}$ were determined after solution of equations (3), solved according to Ritz method. AS co-ordinate functions were selected beam functions, which in case of shell hinge support, look like:

$$U_{nm} = \cos \frac{n\pi\alpha}{\alpha_0} \sin \frac{m\pi\beta}{\beta_0};$$

$$V_{nm} = \sin \frac{n\pi\alpha}{\alpha_0} \cos \frac{m\pi\beta}{\beta_0}; \quad (4)$$

$$W_{nm} = \sin \frac{n\pi\alpha}{\alpha_0} \cos \frac{m\pi\beta}{\beta_0}$$

The optimizing was performed with the help of global search algorithm to within $\varepsilon \approx 2\%$. Calculation results are given in Table 1.

Table 1

Local min	S sm ³	h sm	β_0 rad	σ_i kg/sm ²	Steps
1	125,0179	0,956369	2,467197	1990	52
2	126,6418	1,125607	1,662033	1977	28
3	114,072	0,8343	2,6613	1989	39
4	165,1633	1,53425	1,320312	1993	42

Figure 2 shows curves $\sigma_i(\alpha, \beta)$, corresponding to received minimums.

Task 2. Optimizing of cylinder shell, of rectangular shape, free-supported along edges under uniform load of q intensity.

The shell thickness is variable and is determined according to

$$h = h_0 + h_1 \cdot \sin \frac{\pi\beta}{\beta_0}$$

Optimized parameters: h_0, h_1, β_0 .

$$\frac{\pi}{10} \leq \beta_0 \leq \pi$$

$$0,5\text{cm} \leq h_0 \leq 3\text{cm}$$

$$-0,5\text{cm} \leq h_1 \leq 0,5$$

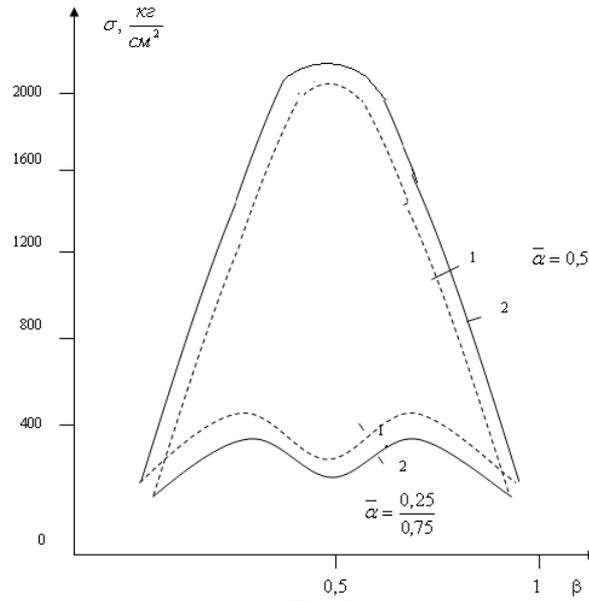


Fig. 2.

The rest parameters are the same as in task 1.

Optimization was performed with the help of global search algorithm to within $\varepsilon \approx 2\%$. Calculation results are shown in table 2.

Table 2

In. loc. min	S sm^2	h_0 sm	h_1 sm	β_0 rad	$\sigma_{i \max}$ kg/sm^2	Steps
1	81,7348	0,38713	0,48340	1,9416	1998,7	57
2	97,8944	0,53173	0,38432	2,3025	1965	24
3	110,0278	1,2699	-0,4740	1,7447	1952	27
4	126,6314	1,0333	-0,1104	2,4939	1946	13
5	114,899	1,0548	-0,21962	2,2839	1936	23
6	106,003	1,0408	-0,0763	1,2513	1999,3	54

Figure 3 shows curves $\sigma_i(\alpha, \beta)$, related to 1-st and 4-th minimums

As we can see from tasks' 1 and 2 solutions results, the usage of cylinder shells of different thickness gives the opportunity to reduce a construction weight by about 14%. Tables 1 and 2 show the effectiveness of global search algorithm use in construction optimization, according to strength criterion.

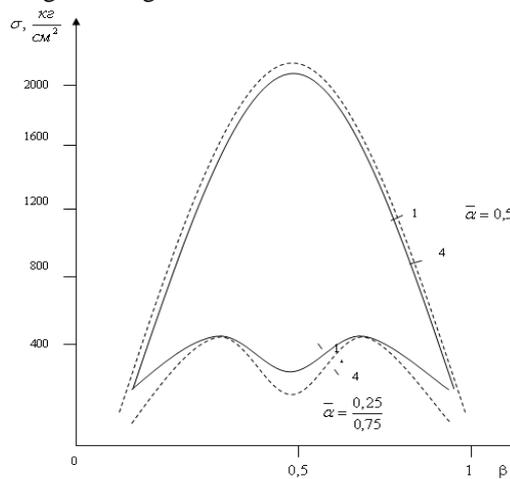


Fig. 2.

If the shell was optimized with the help of global search algorithm with constant step, 22-52 steps were necessary to determine one minimum. The difference of time expenses is essential.

VII. Conclusion

We have studied the specific of engineering construction design process and determined the essence and content of engineering constructions design automation. We have developed engineering constructions optimizing software, based on global search method and worked out the calculation of cylinder shells optimization according to their weight.

Reference

1. Popirin L.S. "Mathematic modeling and energy-heating plants optimization" – M.Energy,1979-p. 410.
2. Nazimova Sh.A., Yakubov S.H. "Algorithm system, automating optimization process for engineering construction design." Ruiz Patent Department Certificate, DGU 01422, 2007.13.11.
3. Kabulov V.K., Nazirov Sh.A., Yakubov S.H. "Optimization tasks solutions algorithms"-Tashkent, "FAN", RU SA, 2008. p. 240.

Authors



He graduated from the Tashkent State University (now National university Uzbekistan) in 1976. In 1976-1991 he worked as an engineer, younger, senior and leading scientific employee; in 1992-2002 and 2008-2011 –as a laboratory manager in the Institute of the Cybernetics of the Sciences Academy of the Republic of Uzbekistan, in 1983 he got the candidate's degree. In 1991 - got a doctor degree. In 2001-2008 he worked as the Head of "Programming Technology" Department, in 2003-2005 – as the Dean of "Information technology faculty. Since 2011 he has been working as the Head of "Informatics Department" in Tashkent Information Technologies University ; he is a doctor of physics-mathematical sciences, professor ; has written more than 400 scientific articles, including 2 monographs and 25 scholastic allowances.



He graduated from the Karshi State pedagogical institute (Uzbekistan) in 1977. In 1977-1981 he worked as the engineer, the assistant. In 1981 to 1987 he studied the postgraduate study in Tomsk State University (Russia) and in 1990 - received a PhD degree on mechanic science. In 1993-2001 he worked as the Head of the chair, in 2001-2005 – as the Dean of the faculty, in 2005-2009 – as the Head of the chair by pulpit, in 2009-2011 – as the Deputy Rector of Karshi State University. Since 2011 he has been working as the Head of Karshi State University Department.

Single Patch E-Shaped Compact Microstrip Antenna

Ahmad Bayat

Department of Electrical Engineering, K.N.Toosi University of technology, Tehran, Iran

Abstract: Micro strip antennas have been the subject of some inventive work in recent years, and are currently one of the most dynamic fields in antenna theory. A single wideband, E-shaped, compact microstrip antenna is presented in this article in order to be employed for high speed WLANs operating in the C band range. Employing only a single patch, a high impedance bandwidth is achieved. The simulated impedance bandwidth ($VSWR < 1.75$) is 30%, and the momentum bandwidth is about 28.5%. The structure of the antenna consists of a perfect conductor on the top of a substrate (RT_DURROID_5880) with a dielectric constant of about 2.2 and a height of 20 mile, which is backed with a perfect conductor ground plane. The impacts of different parameter of antenna are also studied in this article.

Index Terms: E-shaped antenna, ADS2009, RT_Duroid 5880, microstrip

I. INTRODUCTION

In applications of aerospace, satellite and missile which requires small size, easy installation and low cost, antennas with low knob are used like microstrip antennas. Current commercial applications such as mobile communications and wireless telecommunications are taken into account. The printed-circuit antennas use technology simply because they are considered to be made simple and cheap. While installing on the stiff surface, they will have good durability. Appropriate forms can also change diffusion fashion. Microstrip antenna has diversity at the resonant frequency, polarization, pattern and impedance requirements.

The main disadvantages of microstrip antenna can be low efficiency, low throughput, narrow bandwidth and low frequency range. However, it is desirable to use it in some low bandwidth applications. Of course, increasing the thickness of the substrate can expand the bandwidth efficiency. However, this increases the surface waves that is an undesirable factor, because waves can be absorbed in the environment. Surface waves move on the surface layer, scattered in the surface curvature and discontinuities (cuts dielectric surface) and reduce antenna patterns and polarization characteristics.

When bandwidth is large, surface waves can be eliminated using the holes.

Microstrip antennas have reached a stage that can accommodate many other types of antennas for their low power applications.

However, a serious limitation of microstrip antenna is its narrow bandwidth. To strengthen and increase the bandwidth of microstrip antennas, many techniques have been developed. In this paper, by using some of the techniques of traditional microstrip antenna, bandwidth has been increased. The design of triangular patch antennas has been simulated in different frequency cohort. In this paper, using IE3D software, voltage standing wave ratio, end-use, axial ratio, radiation pattern of the designed antenna will be assessed. This paper deals with a single probe feed.

Of the proposed antenna, the impedance bandwidth of 37 percent was achieved.

The increasing use of wireless devices has necessitated the need for larger bandwidths. Hence, the bit rate should be higher or in different standards that are applied in different frequency bands.

Ultra-broadband systems (UWB) as a system with a special focus has been considered because this system can realize the bit rate of up to several hundred Mbps. But the main problem is to design an appropriate antenna system.

Microstrip patch antennas are broadly used due to several advantages they have [1], [5-8], namely the light weight, low volume, low fabrication cost, and their capability of dual and triple frequency operations. Yet, microstrip antennas suffer from some disadvantages too. Narrow bandwidth is a serious one of them and different techniques are used to overcome this limitation, including increasing the thickness of the dielectric substrate, decreasing the dielectric constant [1] and using parasitic patches [2]. These techniques, in turn, have some limitations too, namely the excitation of surface waves and the increase in the antenna size. In wireless local area networks, the antenna employed is in PCMCIA format for which a small size, low volume antenna is requisite. A rectangular patch with a U slot [3] embedded in it will give a broadband antenna, and a single E-shaped patch of 15 mm height provides a bandwidth up to 30%. [11] provided mathematical analysis of microstrip as below:

The electric field radiated from a micro strip antenna provides a boundary between two different dielectrics: air and the substrate material. Due to the slight distortion of the field at the boundary, the patch may appear longer in an electrical sense.

Therefore, we have an effective patch length. There is also an effective relative permittivity when analyzing micro strip antenna. The effective relative permittivity can be defined using the following formula [11]:

E-plane pattern

$$E_{\phi} = \frac{kV_0 w}{2\pi r} e^{-jkr} \left[\sin \theta \left(\frac{\sin \left(\frac{kw}{2} \cos \theta \right)}{\frac{kw}{2} \cos \theta} \right) \right] \quad (1)$$

H-plane pattern

$$H_{\theta} = E_{\phi} / \eta(2)$$

Characteristic impedance of microstrip line feed for $w/h \leq 1$

$$Z_o = \frac{60}{\sqrt{\epsilon_{reff}}} \ln \left[\frac{8h}{w} + \frac{w}{4h} \right] \quad (3)$$

for $w/h \geq 1$

$$Z_o = \frac{120\pi}{\sqrt{\epsilon_{reff}} \left[\frac{w}{h} + 1.393 + .667 \ln \left(\frac{w}{h} + 1.44 \right) \right]} \quad (4)$$

Beam widths E-plane

$$\theta_E \cong 2 \cos^{-1} \sqrt{\frac{7.03\lambda_0^2}{4(3Le^2 + h^2)\pi^2}} \quad (5)$$

H-plane

$$\theta_H \cong 2 \cos^{-1} \sqrt{\frac{1}{2 + kw}} \quad (6)$$

Compared to other methods, transmission line method is the easiest one, which represents the rectangular micro strip antenna as an array of two radiating slots, separated by a low impedance transmission line of certain length.

We can see the fringing at the edges that increases the effective length.

$$\epsilon_{reff} = \frac{(\epsilon_r - 1)}{2} + \frac{(\epsilon_r + 1)}{2} \left(1 + 10 \frac{h}{w} \right)^{-1/2} \quad (7)$$

$$w = \frac{c}{2f_r} \sqrt{\frac{2}{\epsilon_r + 1}} \quad (8)$$

II. E-SHAPED MICROSTRIP PATCH ANTENNA

By incorporating two parallel slots into the rectangular microstrip patch antenna, it becomes an E-shaped microstrip patch antenna, which is simpler in construction. The geometry of this antenna is presented in Figure1.

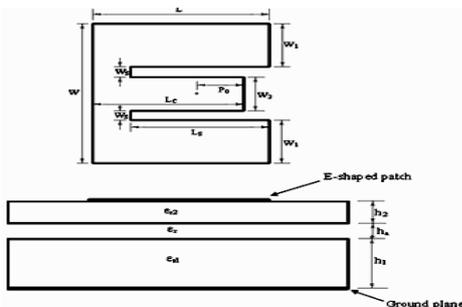


Figure 1. Geometry of the E-shaped microstrip patch antenna

The E-shaped microstrip patch antenna has the width of W_1 , two outer patch strips of the length of L and the width of W , one central patch strip of the length of L_c and the width of W_2 , and also, two slots of the length of L_s and

the width of W_s are introduced symmetrically, regarding the probe position. Employing a coaxial probe, the patch is fed at position P_0 , and two dielectric substrate materials are employed to fabricate the antenna element. An air gap of 1mm is let between the two substrates. The upper substrate is selected to be with high dielectric constant for compact size, and the lower substrate is foam dielectric material in order to provide ground plane. An air gap is let since the thickness of the foam material is fixed, and it helps in optimizing the wide bandwidth.

III. DESIGN OF SINGLE PATCH E-SHAPED MICROSTRIP ANTENNA

We designed the rectangular microstrip patch antenna first, and then incorporated the two parallel slots into the patch. The parameters in designing rectangular microstrip antenna are the followings:

A. Frequency of operation (f_0)

The high-speed computer wireless local area networks function at two frequency bands of (5.15 - 5.35 GHz and 5.725 - 5.850 GHz). The resonant frequency selected for the rectangular patch is the centre frequency of the upper band (5.8 GHz), and with the introduction of the two slots, the E-shaped microstrip antenna covers the lower band.

B. Selection of the substrates

Two substrates are used in fabricating antenna. Upper foam Substrate (RT_DURROID_5880) is of the dielectric constant of $\epsilon_{r1} = 2.2$ and the thickness of $h_1 = 20$ mm. We have let an air gap of 1.0mm between the two substrates. The foam substrate and the air gap are employed for large band width. Lower substrate is of the dielectric constant of $\epsilon_{r2} = 9.6$ and the thickness of $h_2 = 30$ mm.

The microstrip antennas are employed with WLAN adaptor cards in the PCMCIA (also known as PC) format, which have the standard thickness of 5 mm. Therefore, we have limited the total height (4.9624 mm) of the antenna to be less than 5 mm. For designing the microstrip antenna, the advanced Design System (ADS2009) was employed. The dimensions of the optimized antenna element are provided in Table 1.

IV. PARAMETRIC ANALYSIS

The following section deals with the effects of various parameters. By increasing LS , the whole VSWR curve shifts towards lower frequencies. The change is higher in resonant frequency of higher mode, since the relative change in current path length for higher mode is greater than the lower mode current path, as shown in Figure 2. The width WS has a significant effect on the matching to the input port, while it marginally affects the resonant frequencies of the two modes, as presented in Figure 3.

TABLE I. DIMENSION OF OPTIMIZED ANTENNA PARAMETERS

Frequency	5-6 GHz
W	18 mm
L	12 mm
W ₁	5.5mm
L _c	8.5mm
W ₂	5 mm

L_s	8.5mm
W_s	1mm
P_0	2.5mm

Figure 5. Effect of central patch strip length (LC)

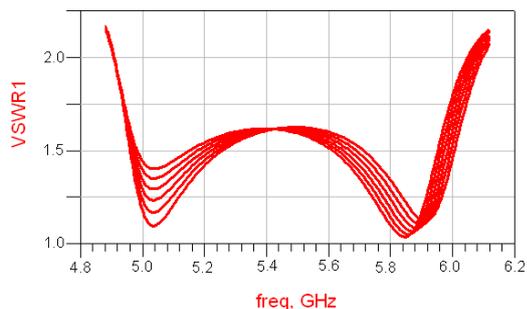


Figure 2. Effect of slot length (LS)

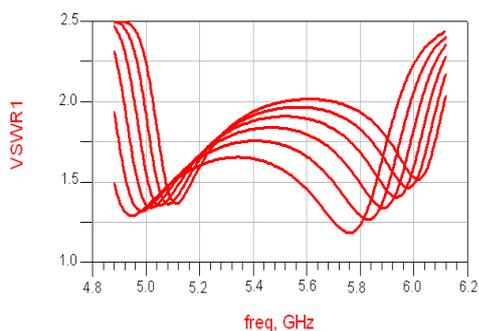


Figure 3. Effect of slot width (WS)

The slot position PS employed for optimizing the matching of both modes is presented in Figure 4.

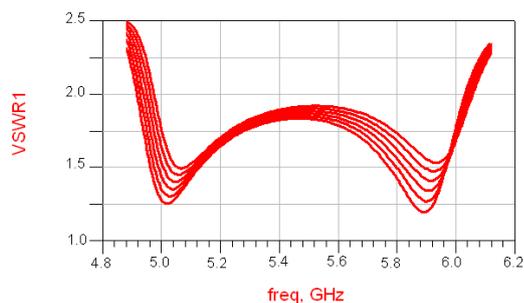
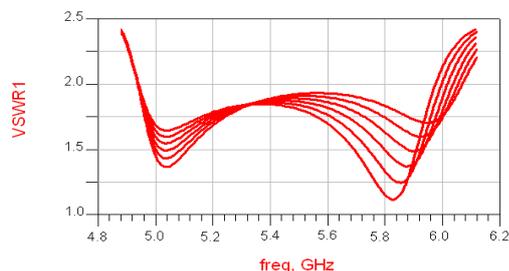


Figure 4. Effect of slot position (PS)

The resonant frequency of the higher mode decreases as L_c increases, while there is no significant change in resonant frequency of the lower mode, as shown in Figure 5.



V. RESULTS

The simulated 2:1 VSWR bandwidth is 30 % covering the (4.82 – 6.43 GHz) frequency band and the momentum 2:1 VSWR bandwidth is 28.8 % covering (4.91 – 6.4 GHz) frequency band, which are presented in Fig. 7. The shift in the frequency band is because of the decrease in the height of the upper substrate with high dielectric constant in the fabrication process. Decreasing the height of upper dielectric substrate is about 0.5 mm, which leads to this shift of frequency band. The radiation pattern was measured in anechoic chamber. Also, there were back lobes in the measured radiation patterns because of the finite size of ground plane. Figure 6 represents Microstrip layout of this antenna.



Figure 6. Microstrip Layout of the E-shaped antenna

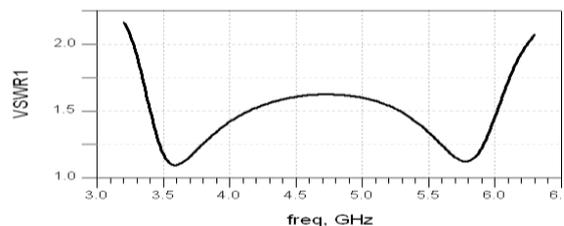


Figure 7. simulated impedance (VSWR < 1.75) bandwidth of wideband E-shaped microstrip patch antenna.

VI. CONCLUSIONS

Regarding the simulated and the measured results, it can be concluded that the E-shaped patch antenna geometry provides wide bandwidth with single patch. The momentum result matches simulated bandwidths. Besides, the antenna shows a good front to back radiation (FBR) ratio. The effect of various parameters of E-shaped patch antenna have been studied without changing the permittivity and the height of the substrates.

REFERENCES

- [1] Ramesh Garg, Prakash Bartia, Inder Bhal and Apsiak Ittipiboon, "Microstrip Antenna Design Hand Book," Artech House, Norwood, MA, 2001.
- [2] D.M.Pozzar "Microstrip Antenna Coupled to Microstripline," Electron Lett., vol. 21, no.2, pp. 49-50, January 1995.
- [3] T.Huynh and K.F.Lee, "Single-layer single patch wideband microstrip antenna," Electron Lett., vol.31, no.16, pp. 1310- 1312, August 1995.

- [4] G. Aiello and G. Rogerson, "Ultra- wideband wireless systems", IEEE microwave Journal, pp. 36-39, June 2003.
- [5] Yi-Lin Chiou, Jin-Wei Wu, Jie-Huang Huang, and Christina," Design of Short Microstrip Leaky-Wave Antenna With Suppressed Back Lobe and Increased Frequency Scanning Region," IEEE trans. Antenna Propagat., vol.57, pp. 3329-3333, Oct. 2010.
- [6] A.Shackelford, K. F. Lee, D. Chatterjee, Y. X. Guo, K. M. Luk, and R. Chair, Smallsize wide bandwidth microstrip patch antennas, in IEEE Antennas and Propagation International Symposium,vol. 1, (Boston, Massachusetts), pp. 86–89, IEEE, July 2001.
- [7] S. C. Gao, L. W. Li, M. S. Leong, and T. S. Yeo, Design and analysis of a novel wideband microstrip antenna, in IEEE Antennas and Propagation International Symposium,vol. 1, (Boston, Massachusetts), pp. 90–93, IEEE, July 2001.
- [8] K. Wong and W. Hsu, —A broadband patch antenna with wide slits, in IEEE Antennas and Propagation International Symposium, vol. 3, (Salt Lake City, Utah),pp. 1414– 1417, IEEE, July 2000.
- [9] Fan Yang, Xui-Xia Zhang, Xioning Ye, and Yahya Rahmat-Sami, "Wideband E-shaped Patch Antenna for Wireless Communications," IEEE Transactions on Antennas and Propagation, vol. 49, no.7, July 2001.
- [10] Constanantine A Balanis, "A ntenna Theory Analsis and Design," John Weily&Sons, New York, 1997.
- [11] Dr Anubhuti Khare, Rajesh Nema, Puran Gour, Rajeev Kumar Thakur "New Multiband E-Shape Microstrip Patch Antenna on RT DUROID 5880 Substrate and RO4003 Substrate for Pervasive Wireless Communication"International Journal of Computer Applications (0975 – 8887), Volume 9– No.8, November 2010

Modeling Of Single-Phase To Three-Phase Drive System Using Two Parallel Single-Phase Rectifiers

Radha Krishna¹, Amar kiran²

^{1,2} (Department of EEE, GIET, JNTU University, Kakinada)

Abstract: This paper proposes a single-phase to three-phase drive system composed of two parallel single-phase rectifiers, a three-phase inverter, and an induction motor. The proposed topology permits to reduce the rectifier switch currents, the harmonic distortion at the input converter side, and presents improvements on the fault tolerance characteristics. Even with the increase in the number of switches, the total energy loss of the proposed system may be lower than that of a conventional one. The model of the system is derived, and it is shown that the reduction of circulating current is an important objective in the system design. A suitable control strategy, including the pulse width modulation technique(PWM), is developed. Experimental results are presented as well.

Index Terms—Ac-dc-ac power converter, drive system, parallel Converter, Fault Identification System(FIS).

I. INTRODUCTION

Several solutions have been proposed when the objective is to supply a three-phase motor from single-phase ac mains [1]-[4]. It is quite common to have only a single phase power grid in residential, commercial, manufacturing, and mainly in rural areas, while the adjustable speed drives may request a three-phase power grid. Single-phase to three-phase ac-dc-ac conversion usually employs a full-bridge topology, which implies in ten power switches. This converter is denoted here as conventional topology.

Parallel converters have been used to improve the power capability, reliability, efficiency, and redundancy. Parallel converter techniques can be employed to improve the performance of active power filters [5]-[6], uninterruptible power supplies (UPS) [7], fault tolerance of doubly fed induction generators, and three-phase drives [8]. Usually the operation of converters in parallel requires a transformer for isolation. However, weight, size, and cost associated with the transformer may make such a solution undesirable [9]. When an isolation transformer is not used, the reduction of circulating currents among different converter stages is an important objective in the system design [10]-[12].

In this paper, a single-phase to three-phase drive system composed of two parallel single-phase rectifiers and a three-phase inverter is proposed. The proposed system is conceived to operate where the single-phase utility grid is the unique option available. Compared to the conventional topology, the proposed system permits: to reduce the rectifier switch currents; the total harmonic distortion (THD) of the grid current with same switching frequency or the switching frequency with same THD of the grid current; and to increase the fault tolerance characteristics. In addition, the losses of the proposed system may be lower than that of the conventional counterpart. The aforementioned benefits justify the initial investment of the proposed system, due to the increase of number of switches.

II. METHODS TO CONNECT SINGLE PHASE TO THREE PHASE DRIVE SYSTEMS

2.1 Static Phase Converter:

Static Phase Converters operate by charging and discharging capacitors to temporarily produce a 3rd phase of power for only a matter of seconds during startup of electric motors, then it will drop out forcing the motor to continue to run on just 1 phase and only part of its windings. Due to their technology, Static Phase Converters do not properly power any class of 3 phase machinery or equipment. They will not in any way power 3 phase welders, 3 phase battery chargers, 3 phase lasers, or any type of machinery with 3 phase circuitry. Static Phase Converters also will not start delta wound 3 phase motors.

2.2 Rotary phase converter:

A rotary phase converter, abbreviated RPC, is an electrical machine that produces three-phase electric power from single-phase electric power. This allows three phase loads to run using generator or utility-supplied single-phase electric power. A rotary phase converter may be built as a motor-generator set. These have the advantage that in isolating the generated three-phase power from the single phase supply and balancing the three-phase output. However, due to weight, cost, and efficiency concerns, most RPCs are not built this way. Rotary Phase Converters Provide Reliable, Balanced, and Efficient Three Phase Power. Quick and Effective Three Phase Electricity.

All converters can be mainly categorized into two groups: one is cascade type and another is unified type [2]. In cascade type, the PWM converter for power factor correction and the PWM inverter for speed control are connected in series with large DC-Link capacitor and two static power converters are operated and controlled in separate. In this type, specific number of switches, to compose the converter and inverter, are required. Therefore, the required number of switches cannot be reduced significantly. On the other hand, in the unified type, conventional concepts of PWM converter and inverter are merged together and same converter handles the functions of PWM converter (power factor correction) and PWM inverter(motor control) at the same time. As an added advantage, the input inductor, which is commonly used in the PWM

converter for power factor correction, can be eliminated and replaced by the existing motor inductor. Therefore, this new concept can significantly reduce the number of components, compared to any conventional cascade type topologies.

III. SYSTEM MODEL

The Conventional system single-phase to three phase system and the Proposed system single phase to three phase systems are labeled as fig 1 and fig 2 respectively as shown below:

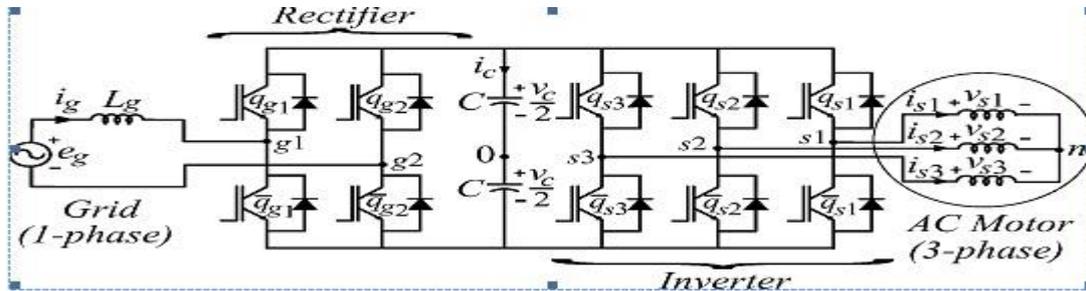


Fig 1 Conventional single-phase to three-phase system

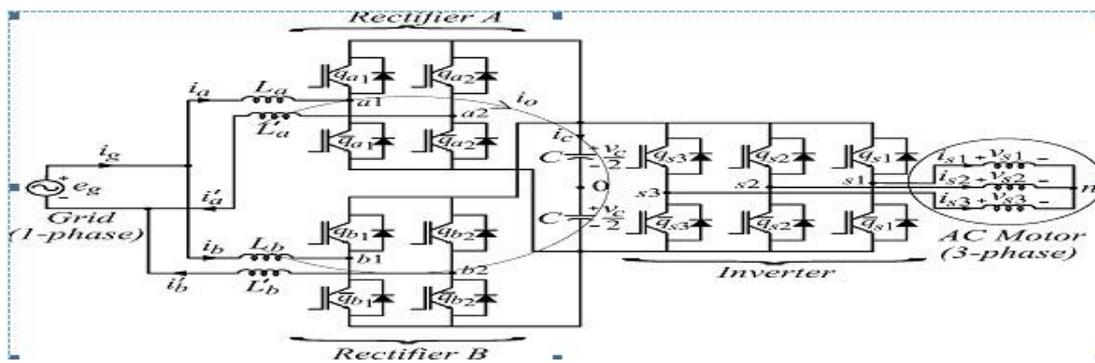


Fig 2 Proposed single-phase to three-phase drive system

The system is composed of grid, input inductors (L_a, L'_a, L_b, L'_b). Rectifiers (A and B), capacitor bank at the dc-link, inverter, and induction machine. Rectifiers A and B are constituted of switches $q_{a1}, \bar{q}_{a1}, q_{a2}$ and \bar{q}_{a2} and $q_{b1}, \bar{q}_{b1}, q_{b2}$ and \bar{q}_{b2} respectively. The inverter is constituted of switches $q_{s1}, \bar{q}_{s1}, q_{s2}, \bar{q}_{s2}, q_{s3}$ and \bar{q}_{s3} .

From Fig. 2, the following equations can be derived for the front-end rectifier

$$v_{a10} - v_{a20} = e_g - (r_a + l_a p)i_a - (r'_a + l'_a p)i'_a \tag{1}$$

$$v_{b10} - v_{b20} = e_g - (r_b + l_b p)i_b - (r'_b + l'_b p)i'_b \tag{2}$$

$$v_{a10} - v_{b10} = (r_b + l_b p)i_b - (r_a + l_a p)i_a \tag{3}$$

$$v_{a20} - v_{b20} = (r'_a + l'_a p)i'_a - (r'_b + l'_b p)i'_b \tag{4}$$

$$i_g = i_a + i_b = i'_a + i'_b \tag{5}$$

Where $p = d/dt$ and symbols like r and l represent the resistances and inductances of the input inductors L_a, L'_a, L_b, L'_b .

The circulating current i_0 can be defined from i_a, i'_a, i_b, i'_b .

$$i_0 = i_a - i'_a = -i_b + i'_b \tag{6}$$

Introducing i_0 and adding (3) and (4), relations (1)–(4) Become

$$v_a = e_g - [(r'_a + r'_a) + (l_a + l'_a)p]i_a + (r'_a + l'_a p)i_0 \quad (7)$$

$$v_b = e_g - [(r'_b + r'_b) + (l_b + l'_b)p]i_b + (r'_b + l'_b p)i_0 \quad (8)$$

$$v_0 = -[(r'_a + r'_b) + (l_b + l'_b)p]i_0 + [(r'_a - r'_a) + (l_a - l'_a)p]i_a + [(r'_b - r'_b) + (l_b + l'_b)p]i_b \quad (9)$$

$$\text{Where } v_a = v_{a10} - v_{a20} \quad (10)$$

$$v_b = v_{b10} - v_{b20} \quad (11)$$

$$v_0 = v_{a10} + v_{a20} - v_{b10} - v_{b20} \quad (12)$$

Relations (7)–(9) and (5) constitute the front-end rectifier dynamic model. Therefore, v_a (rectifier A), v_b (rectifier B), and v_0 (rectifiers A and B) are used to regulate currents i_a, i_b, i_0 respectively. Reference currents i_a^* and i_b^* are chosen equal to $i_g^*/2$ and the reference circulating current i_0^* is chosen equal to 0.

In order to both facilitate the control and share equally current, voltage, and power between the rectifiers, the four inductors should be equal, i.e., $r'_g = r'_a = r'_b$ and $l'_g = l_a = l'_a = l_b = l'_b$. In this case, the model (7)–(9) can be simplified to the model given by

$$v_a + \frac{v_0}{2} = e_g - 2(r'_g + l'_g p)i_a \quad (13)$$

$$v_a - \frac{v_0}{2} = e_g - 2(r'_g + l'_g p)i_b \quad (14)$$

$$v_0 = -2(r'_g + l'_g p)i_0 \quad (15)$$

Additionally, the equations for i_g, i'_a and i'_b can be written as

$$v_{ab} = \frac{v_a + v_b}{2} = e_g - (r'_g + l'_g p)i_g \quad (16)$$

$$v_a - \frac{v_0}{2} = e_g - 2(r'_g + l'_g p)i'_a \quad (17)$$

$$v_a + \frac{v_0}{2} = e_g - 2(r'_g + l'_g p)i'_b \quad (18)$$

In this ideal case (four identical inductors), the circulating current can be reduced to zero imposing.

$$v_0 = v_{a10} + v_{a20} - v_{b10} - v_{b20} = 0. \quad (19)$$

When $i_0 = 0$ ($i_a = i'_a, i_b = i'_b$) the system model (7)–(9) is reduced to

$$v_a = e_g - 2(r'_g + l'_g p)i_a \quad (20)$$

$$v_b = e_g - 2(r'_g + l'_g p)i_b \quad (21)$$

IV. PWM STRATEGY

The PWM strategy for the rectifier will be presented. The rectifier pole voltages $v_{a10}, v_{a20}, v_{b10}$ and v_{b20} depend on the conduction states of the power switches, i.e.,

$$v_{j0} = (2s_{qj} - 1) \frac{v_c}{2}, \text{ for } j = a1 \text{ to } b2 \quad (22)$$

Where v_c is the total dc-link voltage. Considering that v_a^*, v_b^* and v_0^* denote the reference voltages determined by the current controllers.

$$v_a^* = v_{a10}^* - v_{a20}^* \quad (23)$$

$$v_b^* = v_{b10}^* - v_{b20}^* \quad (24)$$

$$v_0^* = v_{a10}^* + v_{a20}^* - v_{b10}^* - v_{b20}^* \quad (25)$$

The gating signals are directly calculated from the reference pole voltages v_{a10}^* , v_{a20}^* , v_{b10}^* and v_{b20}^* . However, (23)–(25) are not sufficient to determine the four pole voltages uniquely from v_a^* , v_b^* and v_0^* . Introducing an auxiliary variable $v_x^* = v_{a20}^*$, that equation plus the three equations (23)–(25) constitute a four independent equations system with four variables (v_{a10}^* , v_{a20}^* , v_{b10}^* and v_{b20}^*). Solving this system of equations, we obtain

$$v_{a10}^* = v_a^* + v_x^* \tag{26}$$

$$v_{a20}^* = v_x^* \tag{27}$$

$$v_{b10}^* = \frac{v_a^*}{2} + \frac{v_b^*}{2} - \frac{v_0^*}{2} + v_x^* \tag{28}$$

$$v_{b20}^* = \frac{v_a^*}{2} - \frac{v_b^*}{2} - \frac{v_0^*}{2} + v_x^* \tag{29}$$

From these equations, it can be seen that, besides v_a^* , v_b^* and v_0^* , the pole voltages depend on also of v_x^* . The limit values of the variable v_x^* can be calculated by taking into account the maximum $v_c^*/2$ and minimum $-v_c^*/2$ value of the pole voltages.

$$v_{xmax}^* = \frac{v_c^*}{2} - v_{max}^* \tag{30}$$

$$v_{xmin}^* = \frac{v_c^*}{2} - v_{min}^* \tag{31}$$

Where v_c^* is the reference dc-link voltages, $v_{max}^* = \max \vartheta$ and $v_{min}^* = \min \vartheta$, with

$$\vartheta = \{ v_a^*, 0, v_a^*/2 + v_b^*/2 - v_0^*/2, v_a^*/2 - v_b^*/2 - v_0^*/2 \}. \text{Introducing a parameter } \mu \text{ (} 0 \leq \mu \leq 1 \text{), the variable } v_x^* \text{ can be}$$

written as

$$v_x^* = \mu v_{xmax}^* + (1 - \mu) v_{xmin}^* \tag{32}$$

When $\mu = 0$, $\mu = 0.5$, and $\mu = 1$ the auxiliary variable v_x^* has the following values $v_x^* = v_{xmin}^*$, $v_x^* = v_{xave}^* = (v_{xmin}^* + v_{xmax}^*)/2$ and $v_x^* = v_{xmax}^*$, respectively. When $v_x^* = v_{xmax}^*$ or $v_x^* = v_{xmin}^*$ a converter leg operates with zero switching frequency. The gating signals are obtained by comparing pole voltages with one (v_{t1}), two (v_{t1} and v_{t2}) or more high-frequency triangular carrier signals. In the case of double-carrier approach, the phase shift of the two triangular carrier signals (v_{t1} and v_{t2}) is 180° . The parameter μ changes the place of the voltage pulses related to v_a and v_b . When $v_x^* = v_{xmin}^*$ ($\mu = 0$) or $v_x^* = v_{xmax}^*$ ($\mu = 1$) are selected, the pulses are placed in the begin or in the end of the half period (T_s) of the triangular carrier signal. On the other hand, when $v_x^* = v_{xave}^*$ the pulses are centered in the half period of the carrier signal.

The change of the position of the voltage pulses leads also to the change in the distribution of the zero instantaneous voltages (i.e., $v_a = 0$ and $v_b = 0$). With $\mu = 0$ or $\mu = 1$ the zero instantaneous voltages are placed at the beginning or at the end of the switching period, respectively, while with $\mu = 0.5$, they are distributed equally at the beginning and at the end of the half period. This is similar to the distribution of the zero-voltage vector in the three-phase inverter.

V. CONTROL STRATEGY

The control block diagram of Fig 2 highlighting the control of the rectifier to control the dc-link voltage and to guarantee the grid power factor close to one. Additionally, the circulating current i_0 in the rectifier of the proposed system needs to be controlled is shown below:

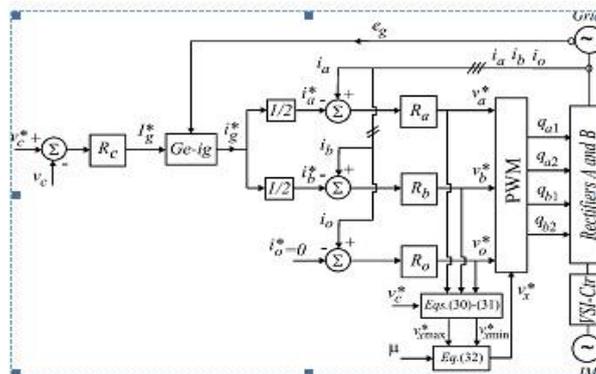


Fig 3 Control block diagram

In this way, the dc-link voltage v_c is adjusted to its reference value v_c^* using the controller R_c . Which is standard Fuzzy logic controllers. This controller provides the amplitude of the reference grid current i_g^* . To control power factor and harmonics in the grid side, the instantaneous reference current i_g^* must be synchronized with voltage e_g , as given in the voltage-oriented control (VOC) for three-phase system. This is obtained via blocks $Ge-i_g$, based on a PLL scheme. The reference currents i_a^* and i_b^* are obtained by making $i_a^* = i_b^* = i_g^*/2$. Which means that each rectifier receives half of the grid

current. The control of the rectifier currents is implemented using the controllers indicated by blocks R_a and R_b . These controllers can be implemented using linear or nonlinear techniques.

The homopolar current is measured (i_0) and compared to its reference ($i_0^* = 0$). The error is the input of Fuzzy controller R_0 , that determines the voltage v_0^* . The calculation of voltage v_0^* is given from (30) to (32) as a function of μ , selected. The motor three phase voltages are supplied from the inverter (VSI). Block VSI-Ctr indicates the inverter and its control. The control system is composed of the PWM command and a torque/flux control strategy.

VI. FAULT COMPENSATION

The proposed system presents redundancy of the rectifier converter, which can be useful in fault-tolerant systems. The proposed system can provide compensation for open-circuit and short-circuit failures occurring in the rectifier or inverter converter devices. The fault compensation is achieved by reconfiguring the power converter topology with the help of isolating devices (fast active fuses— $F_j, j = 1, \dots, 7$) and connecting devices (back to back connected SCRs— t_1, t_2, t_3), as observed in Fig. 10(a) and discussed. These devices are used to redefine the post-fault converter topology, which allows continuous operation of the drive after isolation of the faulty power switches in the converter. Fig. 7.8.1(b) presents the block diagram of the fault diagnosis system. In this figure, the block fault identification system (FIS) detects and locates the faulty switches, defining the leg to be isolated. This control system is based on the analysis of the pole voltage error. The fault detection and identification is carried out in four steps:

1. Measurement of pole voltages (v_{j0}).
2. Computation of the voltage error ϵ_{j0} by comparison of reference voltages and measurements affected in Step (1).
3. Determination as to whether these errors correspond or not to a faulty condition; this can be implemented by the hysteresis detector.
4. Identification of the faulty switches by using ϵ_{j0} .

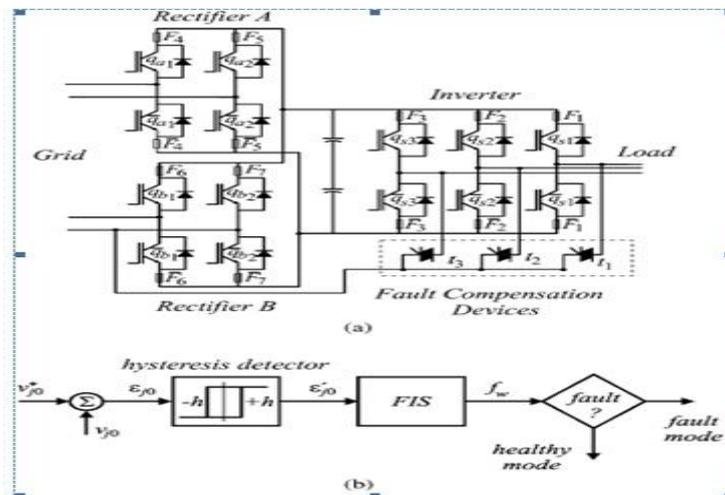


Fig.4 (a) Proposed configuration highlighting devices of fault-tolerant system.
 (b) Block diagram of the fault diagnosis system.

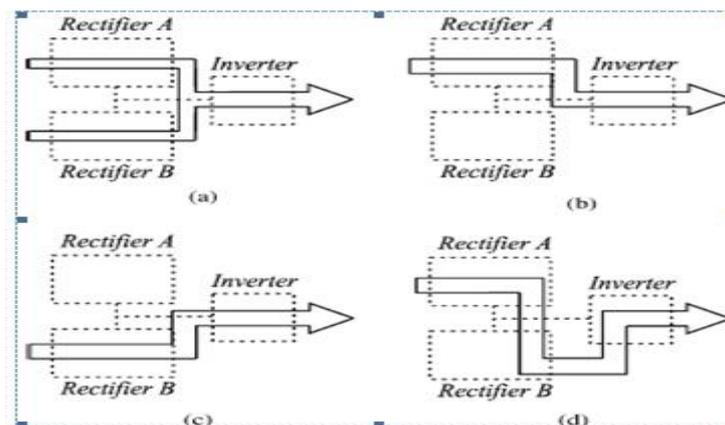


Fig.5 Possibilities of configurations in terms of fault occurrence. (a) Prefault system. (b) Post-fault system with fault at the rectifier B. (c) Post-fault system with fault at the rectifier A. (d) Post-fault system with fault at the inverter.

This way, four possibilities of configurations have been considered in terms of faults:

- 1) pre-fault ("healthy") operation
- 2) post-fault operation with fault at the rectifier B
- 3) post-fault operation with fault at the rectifier A
- 4) post-fault operation with fault at the inverter

Table 1 Efficiency of the proposed system normalized in terms conventional one
 $(\eta_p/\eta_c - 1)$

Frequency/Inductor	S-Ca $\mu=0.5$	D-Ca $\mu=0.5$	D-Ca $\mu=0$
10kHz($i_g=i_g$)	-1.60%	-1.47%	-0.41%
10kHz($i_g=i_g/2$)	3.12%	3.25%	4.36%
5kHz($i_g=i_g$)	-0.74%	-0.27%	1.72%

When the fault occurrence is detected and identified by the control system, the proposed system is reconfigured and becomes similar to that in Fig.1. For instance, if a fault in any switch of rectifier A has been detected by the control system, the whole rectifier needs to be isolated. This isolation procedure depends on the kind of fault detected. If an open-circuit failure is detected, the control system will open all switches of the rectifier A. On the other hand, if a short circuit is detected, the control system will turn on all switches related to rectifier A, and in this case, the fuses will open, and consequently, the rectifier will be isolated, as discussed. Considering now a fault in one leg of inverter, in this case the SCR related with this leg is turned on and the leg $b1$ is isolated, so that the leg $b2$ of rectifier B will operate as the leg of inverter.

VII. MATLAB SIMULINK MODELS

The simulink models of the Proposed converter system, its control strategy and fault diagnosis is also carried out.

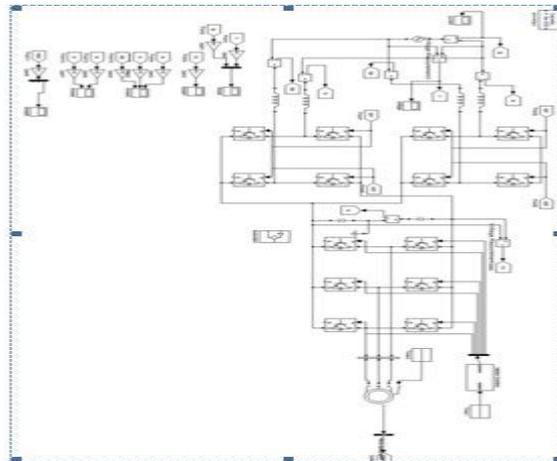


Fig 6 Simulink diagram of single phase to three phase drive system using two parallel single phase rectifiers

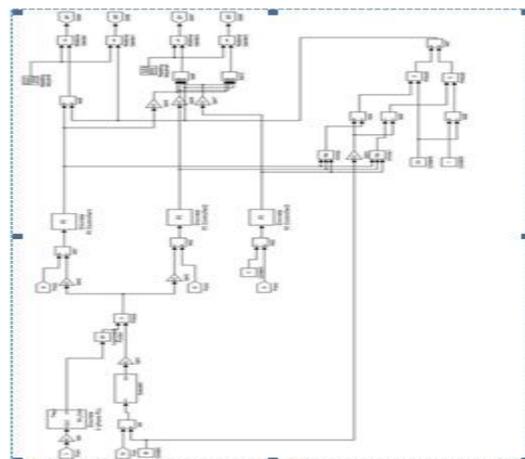
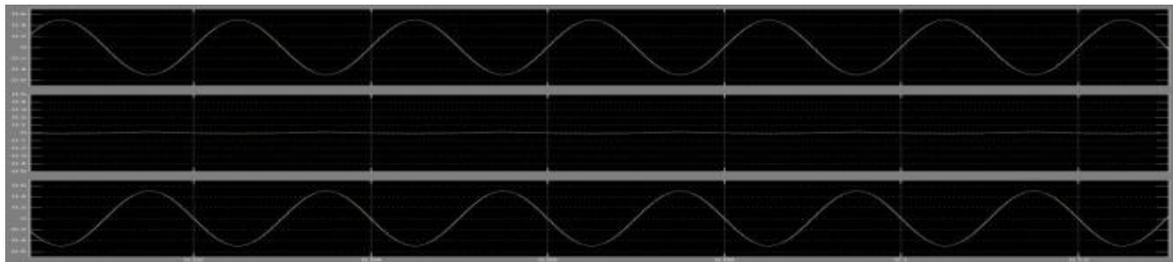
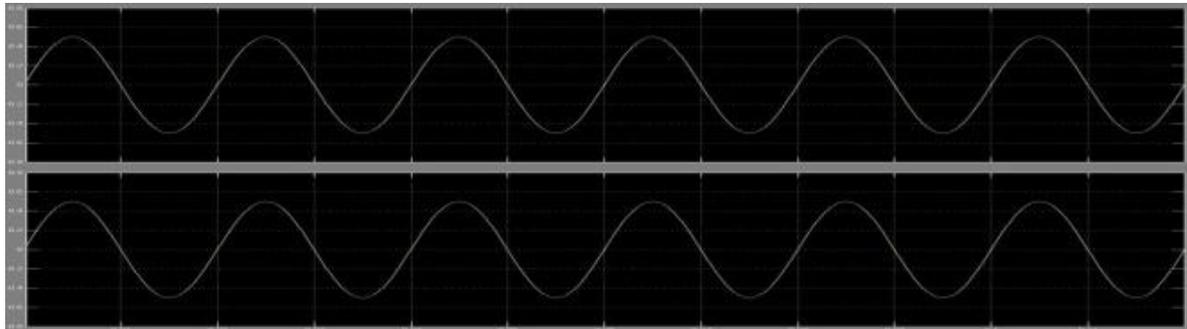


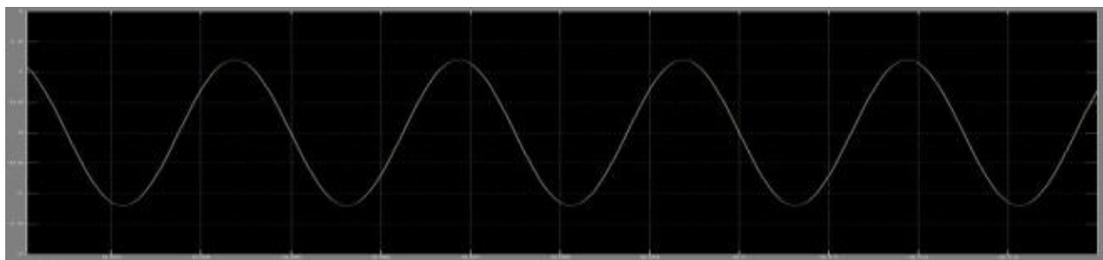
Fig 7 Sub-system



(c)

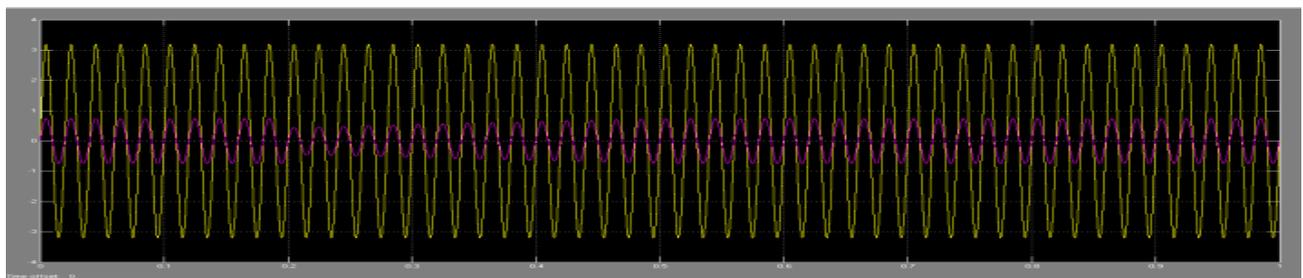


(d)

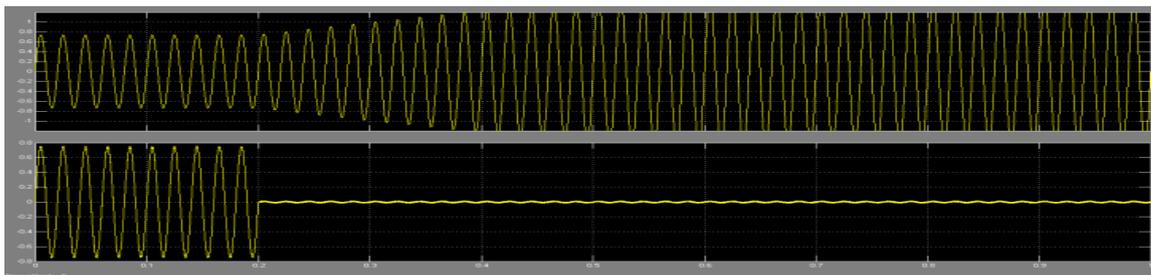


(e)

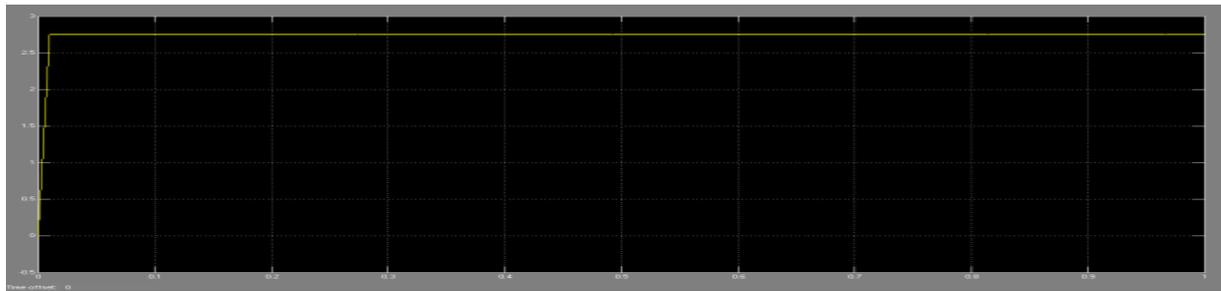
Fig. 10 Shows the waveforms of steady state three phase motor: (a) voltage and current of the grid, (b) dc-link voltage, (c) currents of rectifier A and circulating current, (d) currents of rectifiers A and B, and (e) load line voltage.



(a)



(b)



(C)

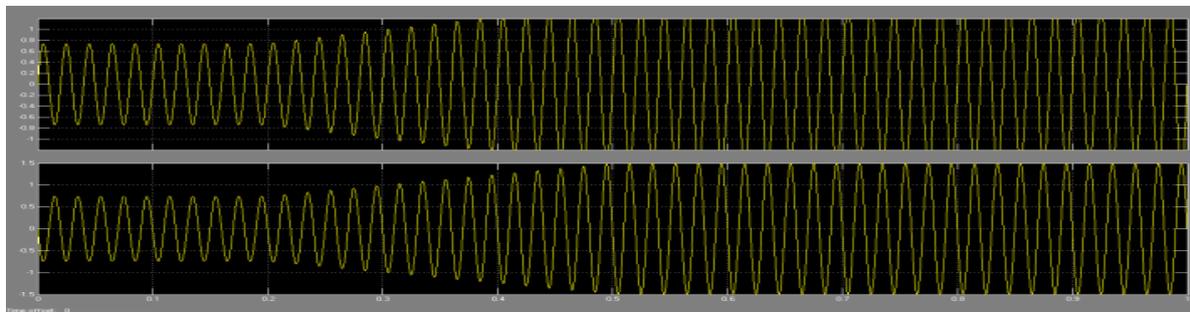


Fig 11 Experimental results when fault occurs at rectifier B. (a) Grid voltage and grid current

(b) Currents of rectifiers A and B (c) Capacitor voltage (d) Currents of rectifier A

IX. CONCLUSION

A single phase to three phase drive system using two parallel converters, three phase inverter and induction motor are proposed. The comparison between conventional and proposed system has been carried out and experimental results are also shown.

REFERENCES

- [1] P. Enjeti and A. Rahman, "A new single phase to three phase converter with active input current shaping for low cost AC motor drives," *IEEE Trans. Ind. Appl.*, vol. 29, no. 2, pp. 806–813, Jul./Aug. 1993.
- [2] B.K. Lee, B. Fahimi, and M. Ehsani, "Overview of reduced parts converter topologies for AC motor drives," in *Proc. IEEE PESC*, 2001, pp. 2019–2024.
- [3] C. B. Jacobina, M. B. de R. Correa, A.M. N. Lima, and E. R. C. da Silva, "AC motor drive systems with a reduced switch count converter," *IEEE Trans. Ind. Appl.*, vol. 39, no. 5, pp. 1333–1342, Sep./Oct. 2003.
- [4] C. B. Jacobina, E. C. dos Santos Jr., E. R. C. da Silva, M. B. R. Correa, A. M. N. Lima, and T. M. Oliveira, "Reduced switch count multiple three-phase machine drive systems," *IEEE Trans. Power Electron.*, vol. 23, no. 2, pp. 966–976, Mar. 2008.
- [5] D.-C. Lee and Y.-S. Kim, "Control of single-phase-to-three-phase AC/DC/AC PWM converters for induction motor drives," *IEEE Trans. Ind. Electron.*, vol. 54, no. 2, pp. 797–804, Apr. 2007.
- [6] L. Woo-Cheol, L. Taeck-Kie, and H. Dong-Seok, "A three-phase parallel active power filter operating with PCC voltage compensation with consideration for an unbalanced load," *IEEE Trans. Power Electron.*, vol. 17, no. 5, pp. 807–814, Sep. 2002.
- [7] L. Asiminoaei, E. Aeloiza, P. N. Enjeti, F. Blaabjerg, and G. Danfoss, "Shunt active-power-filter topology based on parallel interleaved inverters," *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 1175–1189, Mar. 2008.
- [8] M. Ashari, W. L. Keerthipala, and C. V. Nayar, "A single phase parallel connected uninterruptible power supply/demand side management system," *IEEE Trans. Energy Convers.*, vol. 15, no. 1, pp. 97–102, Mar. 2000.
- [9] J.-K. Park, J.-M. Kwon, E.-H. Kim, and B.-H. Kwon, "High-performance transformer less online UPS," *IEEE Trans. Ind. Electron.*, vol. 55, no. 8, pp. 2943–2953, Aug. 2008.
- [10] Z. Ye, D. Boroyevich, J.-Y. Choi, and F. C. Lee, "Control of circulating current in two parallel three-phase boost rectifiers," *IEEE Trans. Power Electron.*, vol. 17, no. 5, pp. 609–615, Sep. 2002.
- [11] S. K. Mazumder, "Continuous and discrete variable-structure controls for parallel three-phase boost rectifier," *IEEE Trans. Ind. Electron.*, vol. 52, no. 2, pp. 340–354, Apr. 2005.
- [12] Z. Ye, P. Jain, and P. Sen, "Circulating current minimization in high frequency AC power distribution architecture with multiple inverter modules operated in parallel," *IEEE Trans. Ind. Electron.*, vol. 54, no. 5, pp. 2673–2687, Oct. 2007.
- [13] J. Holtz, "Pulse width modulation for electronic power conversion," *Proc. IEEE*, vol. 82, no. 8, pp. 1194–1214, Aug. 1994.
- [14] M. P. Kazmierkowski and L. Malesani, "Current control techniques for three-phase voltage-source PWM converters: A survey," *IEEE Trans. Ind. Electron.*, vol. 45, no. 5, pp. 691–703, Oct. 1998.
- [15] B. A. Welchko, T. A. Lipo, T. M. Jahns, and S. E. Schulz, "Fault tolerant three-phase AC motor drive topologies: A comparison of features, cost, and limitations," *IEEE Trans. Power Electron.*, vol. 19, no. 4, pp. 1108–1116, Jul. 2004.

Design and implementation of X band Gunn diode oscillator

Ahmad Bayat

Department of Electrical Engineering, K.N.Toosi University of technology, Tehran, Iran

Abstract: Recently, there has been a revolution in the world of microwave, and Gunn diode oscillators have found many applications and are also heavily applied in research and development in laboratories. These devices are often referred to as an oscillator with low and medium power and are used in microwave receivers. This paper introduces X-band oscillator Gunn diodes and discusses the design and implementation steps that are taken into consideration. The local oscillator of X band, sweeps 8.5 to 9.4 GHz frequency range with a short circuit plate movement across the cavity. Furthermore, change in aperture diameter of diaphragm to achieve maximum output power is examined. We observe that the change in aperture diameter of 8 to 9 mm, leads to 15mW more output power to the amount that can be accounted as outstanding. To prevent damage to the diode, due to current fluctuations of a metal, applying a metal nut with high capacity is also investigated. Phase noise measuring, frequency stability with temperature and time and the second harmonic excitation are also the cases that have been studied. In order to protect Gunn diode in different biases a protection circuit has been designed.

Keywords: Gunn oscillator, frequency mechanical adjustments, protection circuit, isolator, diaphragm aperture.

I. Introduction

Since the late 1940s, in the majority of microwave oscillators and amplifiers, semiconductor devices have replaced light bulbs.

Semiconductor devices are preferred because of their low price and good performance. In the 1960s and 1970s, Gunn and Avalanche diodes were used extensively to produce microwave power [3-1]. Besides, Hybrid and Field-effect transistors, increasingly are used as oscillators, amplifiers, and microwave diodes. However, Impatt and Gunn diodes work at frequencies higher than the Ka frequency band (26-40 GHz) and even beyond it now.

Also in the microwave laboratories, X-band Gunn oscillators are used for low cost and high performance in the educational and research goals.

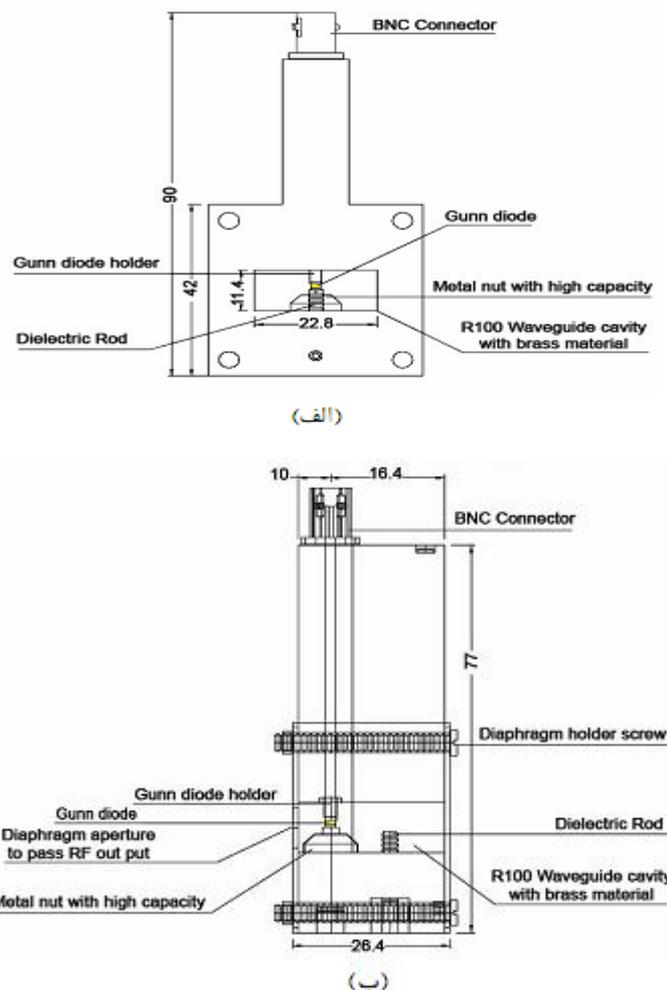
In the cavity, a standard WR100 waveguide and UBR100 flange are used. This oscillator is also used in a wide variety of tests is used as the microwave source. In addition, the Gunn diodes are very fragile in terms of their bias voltage. In order to avoid this problem, an external protection circuit is used.

The resistant circuit will allow the Gunn diodes to be biased in a wide range without any damage. Therefore, if the user reverses the polarity of bias voltage or maximum working voltage of the Gunn diode exceeds, no danger threatens.

II. Gunn oscillator

In order to make this oscillator, a Gunn diode with a maximum flow of 120mA has been applied. The bias voltage is applied through a BNC interface connected directly to the external field. Inside a resonant cavity, in order to optimize its performance, a short circuit plate, a metal plate as a filter and adjustable screws as the passive elements have been used.

By moving short circuit plate along the waveguide, we can change the oscillation frequency of the band, within X band frequencies (8.5-9.4 GHz) as desired. Figure 1 clearly shows these sectors.



a) Longitudinal cut of band X Gunn oscillator
 b) Transverse cut of band X Gunn oscillator

Figure 2 depicts clearly the performance and simple overview of the oscillator. What is also clear in the Figure is the distance between the installation site and

diaphragm aperture of Gunn diode. This distance is equal to half of wavelength which is 16 mm in the built oscillator. Since the relevant frequency of 32 mm-wavelength is approximately 9.4 GHz, the frequency response of the test results is the same as the measured value.

Figure 3 shows a photo of this oscillator.

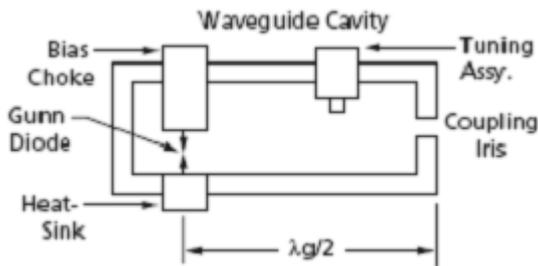


Figure2: The total structure of band X Gunn oscillator [2]



Fig3: A photo of this oscillator

In order to evaluate the performance of the diode in a resonant cavity, we measure the output and examine the oscillator frequency spectrum. It is important to note that for matching the diode with cavity, the measurement of scattering parameters without any bias is required.

Figure 4: shows the frequency spectrum of the oscillator at 9.41387 GHz. It is clear that the received power is about 12 mW.

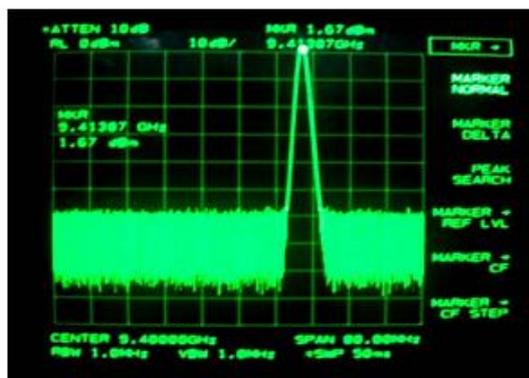
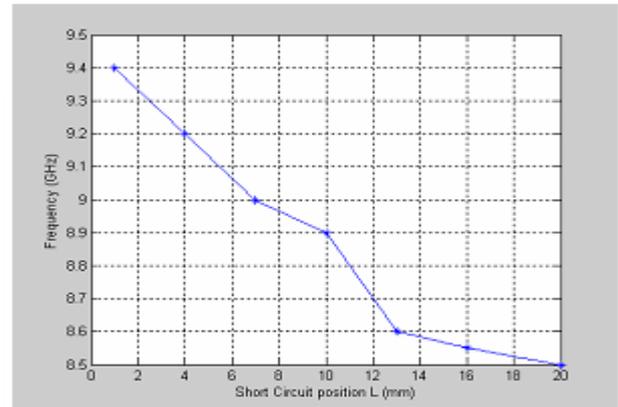


Fig 4: frequency response at 941387 GHz

Also, by connecting a flange into the waveguide mouth and changing the position of short circuit plate

inside it, frequency range of 8.5 to 9.4 GHz is covered. In the following figure this is illustrated.



Frequency changes due to moving the short circuit plate

III. Bias circuit Oscillation

Gunn diode inherently shows negative resistance and this negative resistance with the pre-phase inductance and any excess capacity may lead to oscillation. Permanently, the oscillation is so great that it disables the diode.

To reduce or prevent this type of diode malfunction, high-capacitance capacitors near or parallel to diode will be taken advantage. The combined capacitance must have at least a few tens of MHz frequency response.

IV. Adjustment

Mechanical adjustment of frequency is feasible by using an adjustment rod into the cavity. Rod should be preferably made of a low loss dielectric such as Sapphire with a dielectric constant of 9. Adjustment rate is proportional to the diameter of the rod. If the diameter of the rod is too big, it makes different modes of waves propagated. Off wavelength for a rod with diameter D can be calculated as:

$$\lambda_{co} = 1.7 * D * \sqrt{K} \quad (1)$$

Electronic setup is feasible by using a varactor in the oscillator. Using varactor, regulated bandwidth and oscillator efficiency depend on capacity of connections in chip as well as diode quality coefficient. [4]

V. Protection Circuit

The circuit diagram is shown in Figure 6:

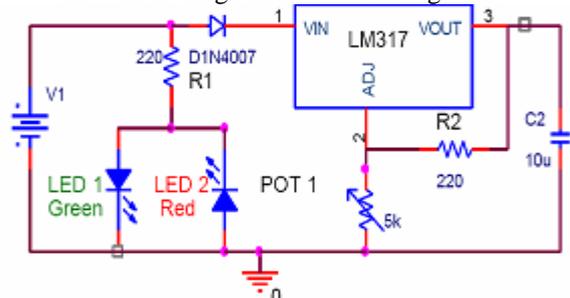


Figure 6: Protection Circuit diagram

Circuit protection is advantageous to biasing Gunn diode in different conditions without causing any harm. If bias voltage polarity gets inversed, the power diode of D1N4007 does not conduct and the output voltage will be zero.

LM317 which is a linear programmable regulator is set on the best bias voltage (14V) for the Gunn Diode. This setting is enabled by changing the potentiometer POT1.

On the other hand, if the input bias voltage exceeds maximum operating voltage of Gunn diode, the linear regulator sets the output voltage on 14 V (or the maximum regulation voltage by POT1). Therefore, even with an increase of DC supply voltage up to 40 V, which is the maximum voltage at the input of the LM317, the Gunn diode will not be harmed.

VI. Frequency stability with temperature

Generally, the oscillator frequency stability depends on cavity material, reactance stability of diode Gunn and varactor. Frequency stability will improve by 3 methods:

- 1-Selecting the appropriate cavity material
- 2- Thermal compensation by incorporating mechanical substances
- 3- Selecting the appropriate Gunn diode

As for cavity material, It is important to say that using an alloy of iron and nickel called Invar is beneficial which is formed by 64% of iron and 46%. The advantage of this material is its thermal expansion which is zero in the range of -50 to 150°C. Aluminum and brass are in the second priority due to having more thermal conductivity, thereby less loss compared to the other materials. Although, the loss of aluminum is less than brass, the brass has been used for four reasons:

- 1-economically inexpensive metal
- 2-Cheap metal in fabrication
- 3-Corrosion resistance
- 4-Easily cut and formed by devices.

VII. Measurement and Laboratory Installation

The module can be seen in figure 7, schematically depicts the RF source. Also, in figure 8, installing structure and measurement result have been shown:

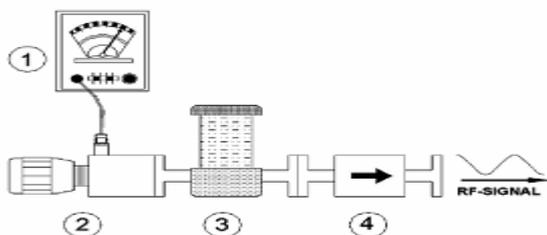


Fig 7: The module Related to RF source and installing and measuring oscillator Gunn

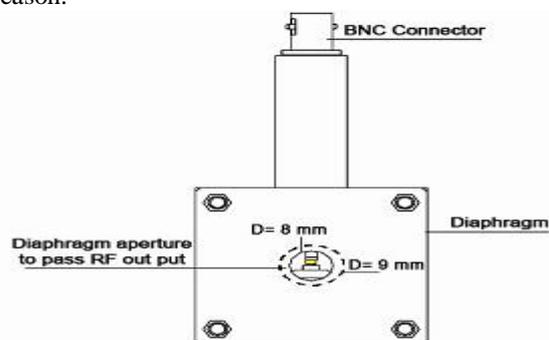


Fig 8: A photo of installing structure and measurement result

In this module a Gunn diode power supply is used to bias the Gunn diode in oscillator. The amount of bias in order to achieve maximum output power is between 8 to 10 volts. The power supply is marked with tag 1. Then, tag 2 is an adjustable cavity of Gunn oscillator. In order to measure the resonant frequency of the cavity, a frequency meter has been used and marked with tag 3. The power reflection from the end of the set is considerable. That is why it may damages the Gunn diode. In order to avoid this issue in section 4, a ferrite isolator is used to make wave passage through one-way and prevent from reflected power passage.

VIII. Output power and diaphragm aperture diameter effect

When the aperture diameter was chosen 8 mm, output power of 12 mW was measured. As it is evident in Figure 9, by changing the aperture diameter to 9 mm, the power increases dramatically and reaches the level of 30 mW. Therefore, we conclude that increasing the aperture diameter increases the output power. However, an excess amount of output power might shorten the life of Gunn diode which is the main part of the oscillator. Passing an excess amount of electrical current through the diode is the reason.



Change in aperture diameter from 8 to 9 mm

IX. Conclusion

In this paper, an X-band Gunn oscillator with WR100 waveguide cavity and the benefits of good performance and low cost have been designed and tested. One of the important points in the design of the oscillator is a metal piece located near the diode that makes a large capacitance. Since the oscillation amplitude is large enough to disrupt the work of the diode, the existence of the metal piece is mandatory in order to avoid any damage. Otherwise, it will not show the desired results. It

should be noted that the frequency response of the combined capacitance must be at least a few tens of MHz. The output power of the oscillator has been evaluated by changes in diaphragm aperture. So, by increasing an aperture diameter from 8 to 9 mm, an increase of 18 mW in output power was observed. The oscillator sweeps the frequency in the range of 8.5 to 9.4 GHz that has many uses. Applications of Gunn oscillator are such as local oscillator in the range of 1 to 100 GHz, microwave amplifiers, radar source, sensors detect the speed, direction, proximity and liquid level, wireless LAN, avoiding colliding vehicles, intelligent traffic control, ABS brake, smart controls of cruise missiles. Also, X band oscillators have been vastly used in research and development especially in labs.

X. Acknowledgement

The author thanks Iran Telecommunication Research Center for its support. Meanwhile, I appreciate Mr Seyyed Mohsen Abu Torab and Mr Ahmad Qaderi for their help.

References:

- [1] S.Hashiguchi and T. Okoshi. "Determination of Equivalent Circuit Parameters Describing Noise from a Gunn Oscillator." 1971 Transactions on Microwave Theory and Techniques 19.8 (Aug.1971 [T-MTT]): pp. 686-691.
- [2] J.F. White. "Simplified Theory for Post Coupling Gunn Diodes to Waveguide." Transactions on Microwave Theory and Techniques 20.6 (Jun. 1972 [T-MTT]): 372-378.
- [3] M.A. Solano, J. Saiz-Ipiña, J.M. Zamanillo and C. Perez-Vega. "X-Band Gunn Diode Oscillator for a Multiple-frequency Continuous Wave Radar for Educational Purposes". IEEE Transactions on Education. Volume 45, No.4, pp 316,322 Nov.2002.
- [4] M.A. Solano, J. Saiz-Ipiña, J.M. Zamanillo and C. Perez-Vega. "X-Band Gunn Diode Oscillator for a Multiple-frequency Continuous Wave Radar for Educational Purposes". IEEE Transactions on Education. Volume 45, No.4, pp 316,322 Nov.2002.

Analysis of Groundwater Quality Improvement Using Rainwater Harvesting: A Case Study of Jamia Millia Islamia

Mohd. Saleem¹, Mugeem Ahmed², Professor Gauhar Mahmood³
Dr. S.A.M. Rizvi⁴

Faculty of Engineering & technology, University Polytechnic, Civil Engineering Section,
Jamia Millia Islamia New Delhi 110025 India

Department of Computer Science Jamia Millia Islamia New Delhi 110025 India

Department of Civil Engineering, Faculty of Engineering & Technology,
Jamia Millia Islamia New Delhi 110025 India

Department of Computer Science Jamia Millia Islamia New Delhi 110025 India

ABSTRACT: *The availability of groundwater neither unlimited nor protected deterioration, in most of the instances the extraction of excessive quantities of ground water has resulted in drying up of wells, damaged ecosystem, land subsidence, saltwater intrusion and depletion of resources. the rate of depletion of ground water level and determination of ground water quality is concern in major cities and towns of countries. It is now recognized that the quality of ground water is as important as its quantity. this paper is an attempt to analyze the impact of rain water harvesting on ground water quality. the area of study is the jamia millia Islamia university new delhi india campus where rain water harvesting structure is installed. therefore the present paper analyses and interprets ground water quality using rainwater harvesting technique. So we can say that the rainwater recharges and improves the quality of ground water which is also depends upon the amount of rain water recharged and the environment of rainwater collection and recharging.*

Keywords: *Intrusion, Deterioration, Depletion, Aquifers*

I. INTRODUCTION

The introduction of the paper should explain the nature of the problem, previous work, purpose, Water is an essential and vital component for our life support system. In tropical regions ground water plays an important role with context to fluctuating and increasing contamination of surface water resources. Ground water has unique features which render it particularly suitable for public water supply. It has excellent natural quality, usually free from pathogens, color and turbidity and can be consumed directly without treatment. Ground water is widely distributed and can be frequently developed incrementally at points near the wafer demand, thus avoiding the need for large water storage, treatment and distribution system. Unfortunately, the availability of ground water is not unlimited nor it is protected from deterioration. In most of the instances, the extraction of excessive quantities of ground water has resulted in drying up of wells, damaged ecosystems, land subsidence, salt water intrusion and depletion of the resources. Ground water quality is being increasingly threatened by agriculture, urban & industrial wastes, which leach or are injected into underlying aquifers. It has been established that once pollution has entered the subsurface environment, it may remain concealed for many years, becoming

dispersed over wide areas of ground water aquifer and rendering ground water supplies unsuitable for consumption and other uses. The rate of depletion of ground water levels and deterioration of ground water quality is of concern in major cities and towns of the country. Being a National Capital Territory, Delhi is facing multifaceted problems regarding water availability, and quality. The population growth and rapid urbanization in and around Delhi has led to immense pressure on basic emanate such as water supply. Deepak Khare et al. (2004) have assessed the impact of RWH on groundwater quality at Indore and Dewas in India using the data from existing tube wells. The roof top rainwater was put through sand filter leading to a reduction in the concentration of pollutants in groundwater. Sharma and Jain (1997) conducted an experiment in Nagpur city where 80,000 liters of water, Collected from the roof top of 100 m² area was recharged. The rise in water level up to 1 m was recorded in the recharge well and adjoining dug wells. The quality of groundwater has also improved as nitrate concentrations got diluted considerably to the desirable limit. Vijaya Kumar (2005) has evaluated the ground water potential by groundwater estimation committee (GEC 1997) norms. Venkateswara Rao (1996) has reviewed the importance of artificial recharge of rainfall for Hyderabad city water supply. A simulation model has been developed by Srivastava (2001) to design a system for determining catchment /command area ratio, size of tank, desirable command area of a single tank and the feasibility/economics of lining a tank. Fayez and Shareef (2009) carried out the research to evaluate the potential for potable water savings by using rainwater in residential sectors of the 12 Jordanian governorates and provided suggestions and recommendations regarding the improvement of both quality and quantity of harvested rainwater. Sazaklia et al (2007) analyzed the quality of harvested rainwater, which is used for domestic and drinking purposes in the northern area of Kefalonia Island in Greece. The influential factors were assessed through a 3-years surveillance. Principal component analysis revealed that microbiological parameters were affected mainly by the cleanness level of catchment areas, while chemical parameters were influenced by the sea proximity and human activities. This paper is an attempt to analyzing the impact of rainwater harvesting on ground water quality of the Jamia Millia Islmia campus which is major challenging issue where the rainwater harvesting structures were installed. The Jamia Millia Islamia is spread over vast area

of 205 acres of land at Jamia Nagar, New Delhi. It has 15% covered area and the remaining area is open space. Depending upon the area of sub-campuses, the campus of the Jamia Millia Islamia is divided into Ten Zones for the installation of rainwater harvesting structures. The rainfall in Delhi area is confined to the month July, August and September (i.e. 90%) and rest of the 10% in the other nine months namely(February, March, April, May, October and November) goes dry. According to Meteorological Department the average rainfall is about 611.0 mm per year with the humidity varying from 17% - 89% and temperature variation from 3° C to 48° C. The evaporation recorded as 1.7 m. The Specific Yield varies in the study area from 16% to 20%.

II. OBJECTIVES

The rainwater harvesting structures were installed at Engineering faculty, Gaddha Colony, department of Fine Arts, University Polytechnic and Administrative Block and the scope of the project is impact of rainwater harvesting on groundwater quality, the steps are to be followed:

1. The following groundwater quality analysis data at pre-installation period of the rainwater harvesting structures at the JMI
2. To collect the groundwater samples in the Post-Monsoon period covering the entire JMI campus area in order to analyses the impact of rainwater harvesting on groundwater quality.
3. Collection of groundwater samples after few months to know the changes in the groundwater quality since the post-monsoon times.
4. Comparative Study of quality of all the three-groundwater quality test results in order to ascertain the overall impact of Rainwater Harvesting on groundwater quality of the campus area.

III. METHODOLOGY

To achieve the main objectives of the study following methodology have been adopted

1. Groundwater Quality Test Results of February 2011 are arranged and analyzed.
2. Groundwater Samples were collected in the month of November, 2011 (Post-Monsoon period) and analyzed.
3. Groundwater Samples were collected in the month of April 2012, in order to know the changes in the groundwater quality since the rainwater recharge.
4. All the three groundwater quality test results of the JMI was compiled and comparative study was done to reveal the overall impact of rainwater harvesting on groundwater quality of the JMI campus.
5. Analysis of rainwater sample and groundwater sample of the JMI Campus in order to check the impact of rainwater harvesting.
6. Study of the impact of rainwater harvesting on groundwater quality potential of the JMI

IV. ADVANTAGES OF ARTIFICIAL RECHARGE

Man plans most artificial recharge projects for specific purpose of saving or storing fresh water for subsequent use. Among these projects some may serve the

dual purpose of eliminating objectionable amounts of water at the land surfaces and, at the same time putting this water into reserve for eventual extraction.

V. HYDRAULIC EFFECTS DUE TO ARTIFICIAL RECHARGE

Two hydraulic effects are generated by artificial recharge as a result of the head, which is applied in the recharge area and the mass of the water, which is introduced into the aquifer through the recharge area, the piezometric effect and the volumetric effect. The piezometric effect results in a rise in the piezometric surface in the unconfined aquifers and /or a rise of the artesian pressure in the confined aquifers. The piezometric effect is related to three main factors. First, it is related to factors which create a damping effect is related to shape of the piezometric surface to the geological and hydraulic boundaries of the aquifer and to the type of location of the recharging device secondly, it is related to quotient T/C (T=transmissivity coefficient; C=replenishment coefficient which is equivalent of storage coefficient). Thirdly, it is related to the artificial recharge yield and the duration of operation. Other factors such as capillary forces water temperature and presence of air bubbles in the aquifers also have in impact on the piezometric effect. The volumetric effect is related to specific yield, replenishment coefficient, the transmissivity coefficient and the boundary coefficient model studies that were checked through filled experiments have demonstrated that the bulk of the recharge water move according to the two systems of flow. One results in a spreading out effect, with a speed related to the recharge flow, the other in the sliding effect, with a speed related to ground water flow.

VI. IMPACT ASSESSMENT OF RAINWATER HARVESTING AT JAMIA MILLIA ISLAMIA CAMPUS

A) Rainwater Harvesting and Groundwater Quality

The rainwater harvesting is done primarily for the qualitative improvement, irrespective of its methods, whether collecting rainwater in ponds or reservoirs for future use or by recharging the rainwater to the groundwater aquifers through bore hole drilled for the purpose. Apart from the qualitative improvement of groundwater as a result of dilution of certain chemical constituents and dissolved solids. This qualitative improvement of groundwater is utmost importance because where there is saline groundwater or the chemical constituents are more than the desirable or maximum permissible limits, the rainwater recharging the aquifers dilutes it to make it useful for drinking and other proposes, very often. The urbanization, agricultural development, and discharges of municipal and industrial residues into the water resources significantly alter its characteristics. The prevailing climatic conditions, topography, geological formations and use and abuse of this vital resource have significant affect on the characteristic of the water, because of which its quality varies with locations. The term water quality criteria may be defined as the "Scientific data evaluated to derive recommendations for characteristics of water for specific use". The quality analysis conducted on

the groundwater samples collected during the month of February 2011, in the post-monsoon period in the month of November 2011 and in April 2012 reveals this truth of qualitative improvement of groundwater through rainwater recharge to the aquifers,

B) Sampling and Analyses of Groundwater

The ground water quality monitoring studies were undertaken during the months of February 2011, November 2011 and April 2012. As the water systems are heterogeneous to varying degrees in space and time, the water samples were collected at from Tube wells and injection wells located at Engineering faculty, Gaddha Colony, Department of Fine Arts, University Polytechnic and Administrative Block, covering the entire campus area. All the ground water samples were preserved in the field itself and transported to the laboratory for water quality testing.

C) Qualitative Impact of Rainwater Harvesting, at JMI

The groundwater quality of Jamia Millia Islamia has been tested in the month of February 2011 by collecting 10 groundwater samples from the existing tube wells. The following table-1 shows the parameters considered for the test and their results:

TABLE-1: GROUNDWATER QUALITY ANALYSIS (FEBRUARY, 2011)

S. No.	Parameters	Values
1.	pH	7.0 – 8.1
2.	Total Hardness	240 – 778 mg/l
3.	Nitrate	4.1 – 145.6 mg/l
4.	Chloride	96 – 828 mg/l
5.	Sulphate	70 – 270 mg/l
6.	Flouride	0.2 – 1.4 mg/l
7.	TDS	360 – 1337 mg/l

The concentrations of various parameters at the entire Jamia campus are shown in the above table 1. Later on, groundwater samples were taken twice during the months of November 2011 (13 samples) and April 2012 (13 samples) from the existing tube wells and injection wells at the campus for the comparative analysis of the groundwater quality before and after rainwater recharging

The comparative analysis of the groundwater quality before and after rainwater recharging.

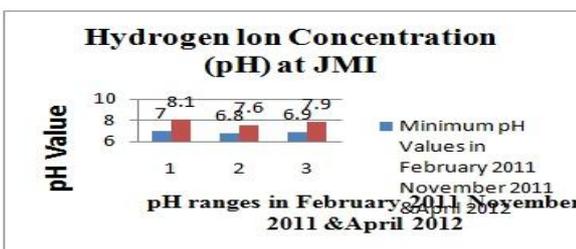


Figure: 1 Comparative Hydrogen Ion concentrations (pH)

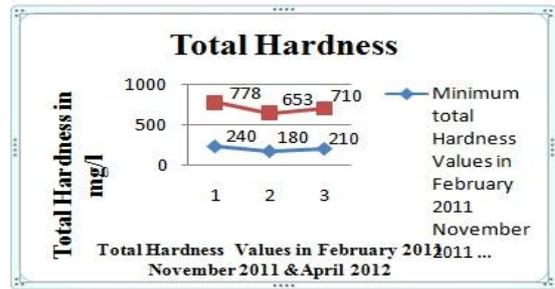


Figure: 2 Comparative Total Hardness

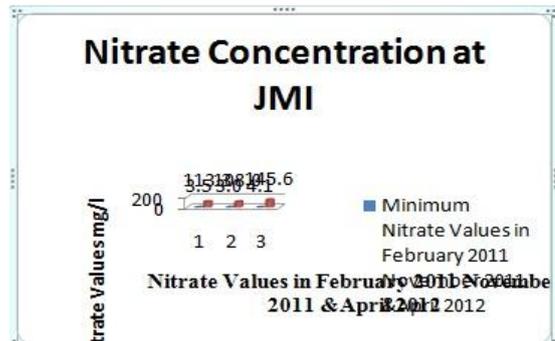


Figure: 3 Comparative Nitrate Concentrations

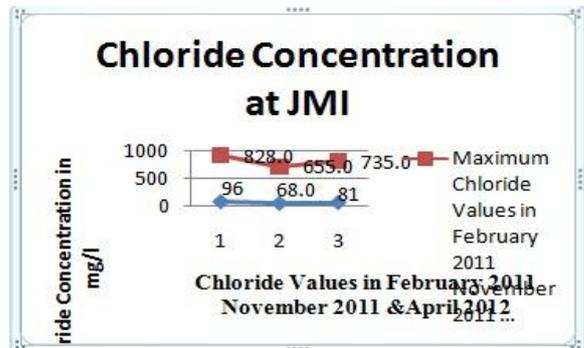


Figure: 4 Comparative chloride Concentration

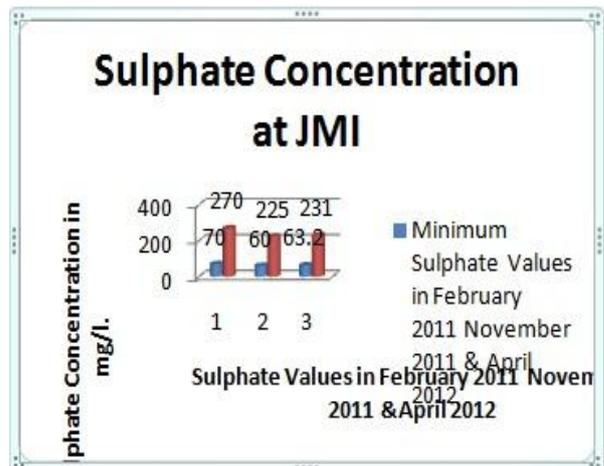


Figure: 5 Comparative Sulphate Concentration

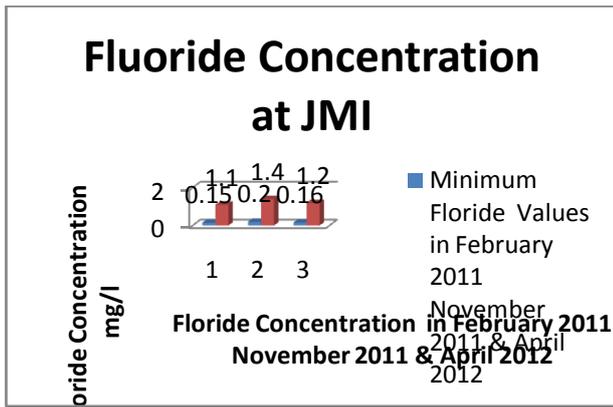


Figure: 6 Comparative Fluoride Concentration

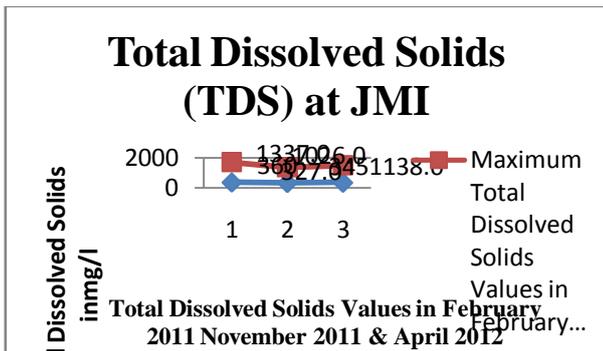


Figure: 7 Comparative Total Dissolved solids (TDS) Concentration

From the account of analysis of groundwater quality it is evident that the significant reduction in concentration of different chemical constituents is the result of dilution of these substances by the rainwater, which is pure and has the pH 5.2. Again, these chemical constituents attain slightly higher concentration level due to the withdrawal of groundwater and the dispersal of the localized recharged water to the downstream areas

D) Rainfall during the Year 2011 and Impact on Groundwater Quality

The south-west monsoon of 2011 was marked by near-normal rainfall over the country distributed equitably over both space and time. In Delhi, the rainfall recorded during 2011 was 748 mm, higher than the average annual rainfall of about 611.0 mm. This 137.0 mm increase in rainfall resulted in the greater than the normal recharge, The Rainwater Potential of the campus area, i.e., amount of rainwater recharge to the aquifer during the year 2011 was 49 % more than the average potential. Hence, it had a bearing on the groundwater quality. As a result, the pH and other chemical constituent have shown the substantial decrease in their concentration by the dilution through the rainwater and after a few months (in April 2012), the slight increase was recorded in the concentration as depicted in the comparative groundwater quality test data in the table-2.

Table-2 Comparative Groundwater Quality Analysis

S.	Constituents	Indian Standard	Range of water quality parameter tested
----	--------------	-----------------	---

No.	Parameter	Desirable limits	Maximum Permissible limits	February 2011	November 2011	April 2012
1	pH	6.5-8.5	6.5-9.5	7.0-8.1	6.8-7.6	7.1
2	Total Hardness (mg/L)	300	600	240-778	180-653	360
3	Nitrate (NO ₃) (mg/L)	45	100	4.1-145.6	3-108	1337
4	Chloride (mg/L)	250	1000	96-228	68-655	1026
5	Sulphate (SO ₄) (mg/l)	200	400	70-270	60-225	1337
6	Fluoride (F) (mg/L)	1.0	1.5	0.2-1.4	0.15-1.1	0.16
7	Total Dissolved Solid (mg/L)	500	2000	360-1337	327-1026	1026

The groundwater quality test result for each of the samples collected during the months of February 2011, November 2011 and April 2012

E) Salinity Scenario

The Jamia Millia Islamia campus has both the fresh and saline groundwater extractions. The saline water availability is more than fresh. The salinity increases with depth like in Mujeeb bagh area. The rainwater recharged to the aquifer dilutes the salt concentration in ground water.

VII. CONCLUSION

The rainwater harvesting structures were installed at various at the JMI at the 20 meters depth. The other dimensions of ponds and injection wells are varying as per the designs and cross-sections enclosed. It is said while going for the installation of rainwater harvesting structures at JMI that the declining trend of the groundwater level will occur if the rainwater harvesting system will not be adopted. There was the mention of the improvement in groundwater quality of the area. this paper is an effort to analysis the impact of rainwater harvesting at the JMI, qualitatively. The ground water samples collected analyzed in the months of February 2011, November 2011 and April 2012. These data were analyzed and compared to know the analyses reveals a very interesting figure as for as the ground water quality improvement is concerned and approves the claim of quality improvement.

Therefore, in the light of the present work, analysis and interpretations, it can be concluded that the rainwater recharge improves the quality of groundwater and its quality depends upon the amount of rainwater recharged and the environment of rainwater collection and recharging

REFERENCES

- [1] A.P.H.A., Standard methods of analysis of water and waste water, *American Public Association U.S.A. 19th edition. 1995*
- [2] Aml K. & Amas Heavy metals removal from water using phy. Ash Simanta Falli, santiketan (Distt Birbhum, W.Bengal).
- [3] BIS, Indian Standard specification for drinking water IS: 10500, *Indian Standard Institution, New Delhi. 1983,*
- [4] BIS, , Guidelines for the quality of irrigation water, *Indian Standard institution, New Delhi. 1986*
- [5] Central Ground Status of groundwater quality including pollution aspects water Board, , *India, Central Ground Water Board, Ministry of Water Recourses, Govt. of India, New Delhi. 1997*
- [6] Central Ground Treatment techniques for safe drinking water, Central Water Board, *Ground Water Board, Ministry of water Recourses, Govt. of India New Delhi. 1998*
- [7] Ground water quality series GWQS/4, Ground water quality in problem areas. *A status report (Part-IV), CPCB, Delhi, , 1995*
- [8] Hand B.K. Technical report on hydro chemistry and water pollution, *C.G.W.B. Ministry of Irrigation. 1981*
- [9] Hem J.D., study and interpretation of chemical characteristics of natural waters *U.S.G.S., WSP 1473p.p.-1-338 1970*
- [10] I.C.M.R., , Manual of standards of quality for drinking water, *Indian Council of Medical Research, New Delhi. 1975*
- [11] Kakar Y.P., Sikka V.M. Janeshwar, Das. Hydro-chemical and pollution of ground water in Faridabad area (Haryana), *Central ground Water Board (Govt. of India). 1989*
- [12] Minhas P.S.Gupta, R.K. Quality of irrigation water assessment and management published by *Indian Council of agricultural Research, New Delhi.. 1992*
- [13] Programme objective series Problem/34., Ground water quality I n the union territory of Delhi (Abridged report) *Central Board for the prevention and control water pollution, New Delhi. 1985-86*
- [14] Series NAAQM/5 Central Pollution Control Board (Ministry of Environment & Forest, Govt. of India), *Privesh bhawan.'east Arjun Nagar, New Delhi. / 1995-96.*
- [15] Turney G.L.. Quality of ground water in south eastern and south-central, *Washington, U.S, Geological survey, Water resources Investigatory report SH-4262 Washington. 1986*
- [16] US, *US Federal Water Control Administration. 1968.*
- [17] WHO, *Guidelines for drinking water second edition Vol. 1&2, World Health organization, Geneva. , 1996*
- [18] Ala Eldin M.E.H. Sami Ahmad M. Guruadha Rao, Dhar RI. Groundwater modeling in part of Central Ganga Basin, *Uttar Pradesh, India. Technical Report, No AMUINGRI I, Collaborative Project between Remote Sensing Application for Evaluation and Geo-Engineering, Aligarh and NGRI. 1998*
- [19] Baweja, B. K. and Karanth, K. R. (1980) *Groundwater Recharge Estimation in India, Tech. Sr.H. Bull 2, Central Groundwater Board*
- [20] Bear. J, *Hydraulics of Groundwater McGraw Hili, New York: 569 pp. 1979*
- [21] Copen. T.B. *Uses of Environmental Isotopes, in Regional Ground Water Quality 1993*
- [22] Dincer. T. Davis, *C.HApplication of Environmental isotope tracers to modeling in hydrology 68: 95 pp .1984*
- [23] Gooper, Pf, Hubson J. A & Joens: Sewage Treatment by Reed Bed Systems, *Journal of the Institute of Water & Environment Management.*

Comparative study of prestressed steel – concrete composite bridge of different span length and girder spacing

Vikash Khatri¹, Pramod Kumar Singh² and P.R.Maiti³

Department of Civil Engineering, Indian Institute of Technology, Banaras Hindu University, Varanasi-221005, India

ABSTRACT: Prestressed Concrete and Steel Concrete Composite (SCC) are commonly used for constructing bridges. Construction of prestressed concrete is time taking and lower reliable. SCC bridges have problem of excessive deflection under dead and live load, and deflection due to shrinkage and creep of deck slab concrete.

External post-tensioning for strengthening of existing bridges has been used in many countries and has been found to provide an efficient and economic solution for a wide range of bridge types and conditions. External prestressing is now being used for construction of new bridges also.

This paper introduces a new concept of Prestressed Steel-Concrete Composite (PSCC) bridge, in which external post-tensioning is used in the SCC bridge. In the PSCC bridge, high tensile wires are tensioned by means of jacks bearing on the end block of the concrete deck slab and anchored. As a result, longitudinal stress level of the concrete deck slab is raised, which not only eliminates shrinkage and creep strains but also improves its fatigue performance.

In the present study effects of the total area of steel girder, prestressing force required in the cables, and stress in the deck slab are presented for various span lengths and girder spacings. The total steel girder area required in 4-girder system is nearly 20% lower than that of 5-girder system. Stresses in the deck slab due to prestressing were raised between 2 N/mm² to 10 N/mm² for 4-girder system. In the 50% of live load hogging deck case, the range of stresses in deck slab is lower than that of the no hogging case. Maximum stress in the deck slab for 4-girder system with the 50% of live load hogging case is also reduces to 9.96 N/mm² from 12.27 N/mm² in comparison to the no hogging case. It is concluded that prestressing raises stress level of the deck slab concrete resulting in its better fatigue performance, and also improves strength and stiffness of the bridge considerably.

Keywords: Prestress, Composite, Bridge, Shrinkage, Post-tension

I. INTRODUCTION

For highway bridges, composite bridge deck using a reinforced concrete slab over two or more steel girders is more popular than steel bridge and concrete bridge. The composite bridge also reduces noise and vibration levels in comparison to steel bridge, and is, therefore, environmentally friendly. Composite bridges are lighter, guarantee better quality and have easier and faster erection than concrete bridges [7, 13].

In recent trend pre-stressed concrete bridges [4, 14] have been expanding the applicable span length and are becoming a hard competitor against steel bridges and concrete bridges. Steel bridges, therefore, need new ideas to

regain competitiveness. Steel plates have high tensile strength but are relatively vulnerable to buckling caused by compressive forces and need to be stiffened and strengthened. Resistance against buckling of the composite structure increases when steel girders are combined with reinforced concrete deck slab.

The use of external post-tensioning for the strengthening of existing bridges has been reported [1, 5] to provide an efficient and economic solution for a wide range of bridges. The technique is growing in popularity because of the speed of construction and hence external post-tensioning for prestressed concrete bridges is used in many countries in the construction of new bridges [15]. While new bridges are constructed using external post-tensioning, over the last two decades external post-tensioning has also been considered as one of the most powerful techniques for structural strengthening and rehabilitation.

The presence of early age transverse cracking in concrete bridge decks is often what leads to the eventual structural deficiency of bridges in the long run [2, 17], because these cracks permit the ingress of harmful substances into concrete bridge decks. With the presence of cracks in concrete bridge decks, water, sulfates, chlorides, and other potentially corrosive agents able to permeate to the interior of the bridge deck and cause further deterioration in the form of even larger cracks, spalling, potholes and eventually a loss of cross section of the bridge deck or reinforcing steel, which ultimately leads to an unsafe bridge. The repair of concrete bridge decks is often difficult and expensive because alternate routes are sometimes difficult or impossible to come by. To prevent deterioration from starting in the first place, concrete must not be allowed to crack, especially at an early age. For a concrete structure, to be serviceable, cracking must be controlled and deflections must not be excessive. It must also not vibrate excessively. Concrete shrinkage plays a major role in each of these aspects of the service load behaviour of concrete structures [16, 19].

Serviceability failures of concrete structures involving excessive cracking and/or excessive deflection are relatively common. Numerous cases have been reported, in Australia, Europe and elsewhere [6, 7], of structures that complied with code requirements but still deflected or cracked excessively. In a large majority of these failures, shrinkage of concrete is primarily responsible. Clearly, the serviceability provisions embodied in different codes [8, 9] do not adequately model the in-service behaviour of structures and, in particular, fail to account adequately for shrinkage.

Cracking can significantly reduce the service life of concrete bridge decks, pavements and other concrete structures. If cracking due to concrete shrinkage could be

eliminated, a bridge's service life could be two to three times longer, and costly repairs could be avoided [3, 18].

In the prestressing technique [10], the prestressing tendons are placed outside the concrete section and the prestressing force is transferred to the concrete by means of end anchorages, deviators and saddles.

In the PSCC bridge, the external tendons may be used for prestressing of the composite reinforced concrete deck slab with steel plate girder for long span bridges. The reinforced concrete deck slab is first cast over the plate girder incorporating end anchorages and ducts to house the tendons. When concrete attains sufficient strength, the high tensile wires are tensioned by means of jacks bearing on the end block of the deck slab and anchored. Singh [20] explained that the prestressing force is transmitted to the deck slab concrete which will raise the stress level of the

concrete deck slab and improve its behaviour under fatigue loading significantly, and eliminate the deck slab concrete shrinkage. Additional undesirable compression due to anchoring of tendons in the plate girder is also avoided by anchoring the tendons in the end block of the deck slab concrete.

The primary objective of this paper is to compare the total area of steel girder and prestressing force required in the cables, and stresses in the deck slab using various span lengths and girder spacing.

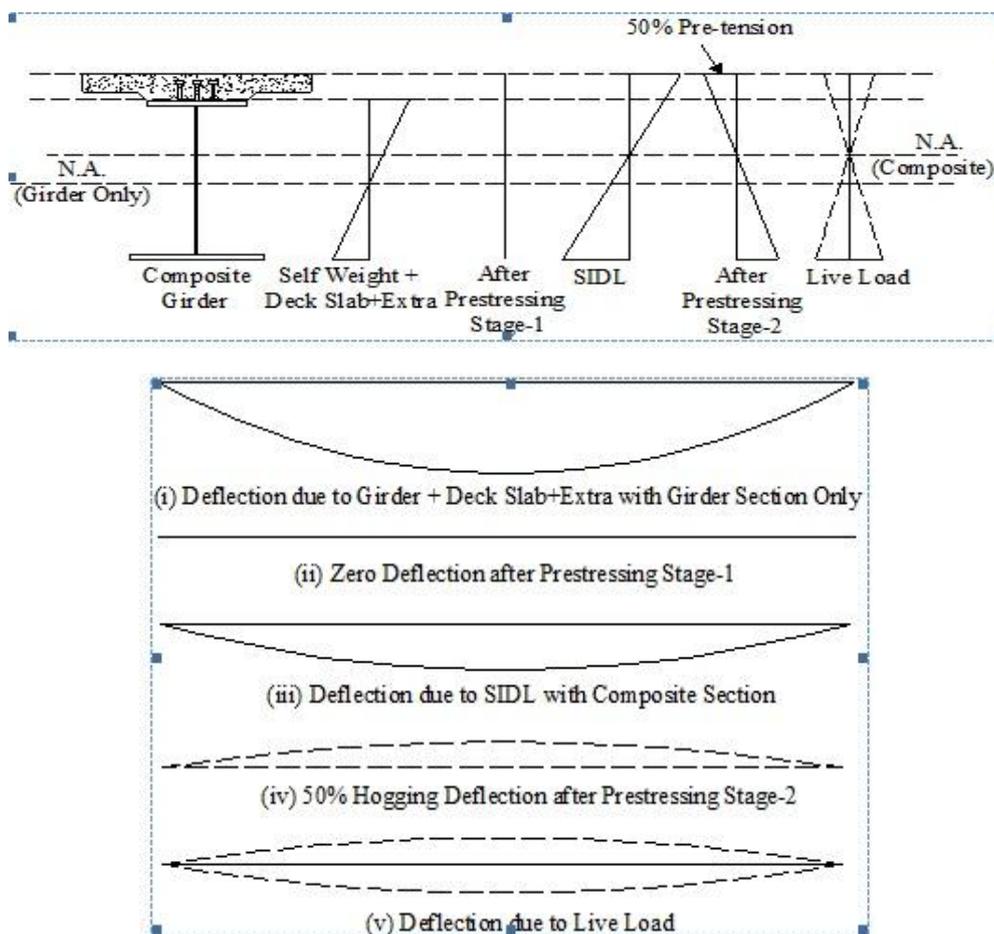


Fig. 1. Variation of flexural stress and deflection for 50% of LL hogging case

Parameters which are common for all the span bridges are:

Grade of concrete	=	M40
Width of deck	=	12 m
Thickness of deck	=	200 mm
Number of cross beams	=	7
Number of cables	=	14

Position of cables = Anchored at the bottom end of cross beams

The following variable parameters were considered in this study:

- Span lengths: 20 m, 40 m, 60 m, 80 m and 100 m
- Girder spacing: 4 girders at 3.0 m and 5 girders at 2.5 m
- Hogging case: 50% of live load deflection

In the hogging case (Fig. 1), in PSCC bridges prestressing is done intentionally to develop upward deflection (hogging) so that bridge curvature under live load is reduces to half. Prestressing is done to provide initial hogging deflection of the 50% of full LL. In service condition the deflection and flexural stresses from the horizontal references are only the 50% of full LL deflection and flexural stresses of the bridge.

II. ANALYSIS OF BRIDGES

PSCC Bridges are designed as per Indian Standards [11, 12] for comparison. Figure. 2 show the typical design of 40.0 m span bridge with 5-girder system.

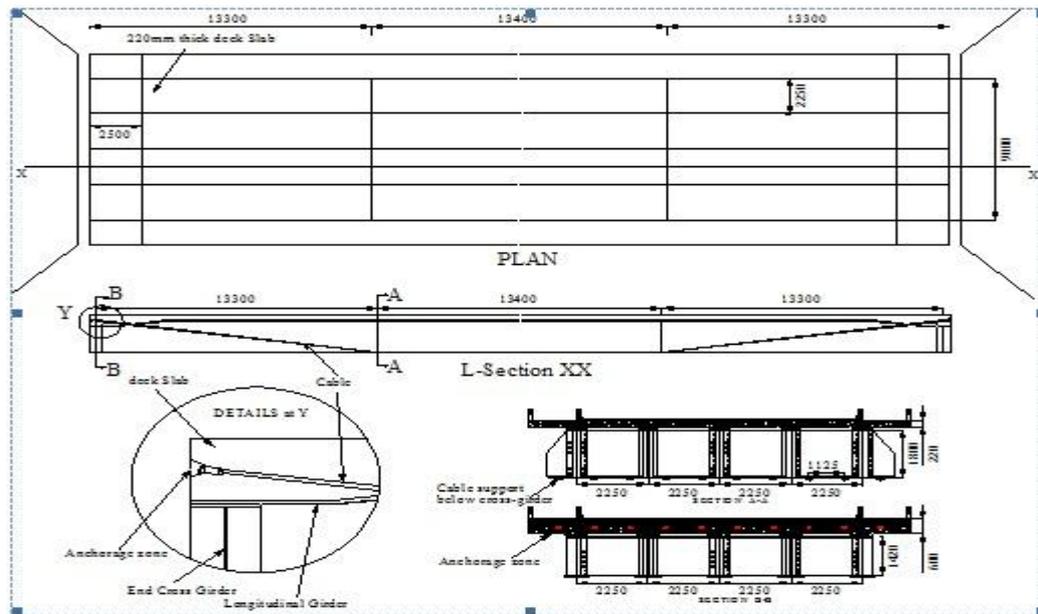


Fig.2. Typical design of 40.0 m span PSCC Bridge with 5-girder system

Figures 3-a and 3-b shows the cross section of 5-girders and 4-girders system, respectively.

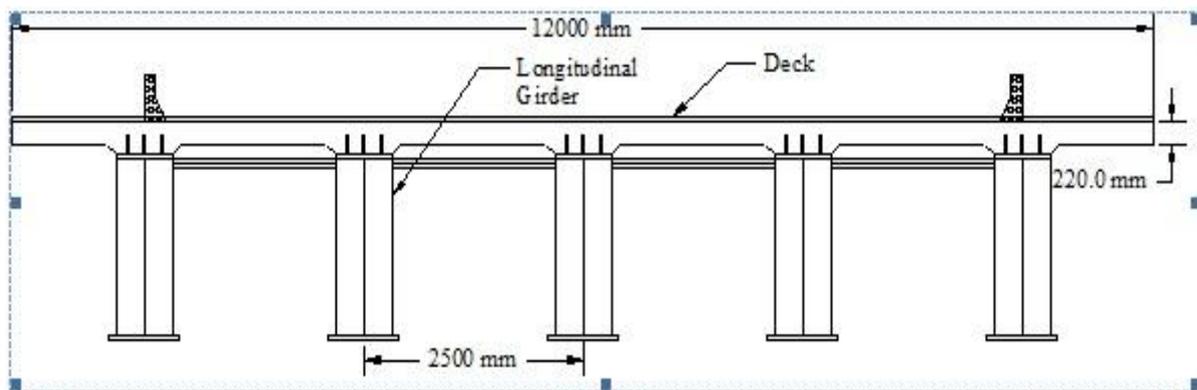


Fig. 3.a Cross Section of 5-Girders Composite Bridge

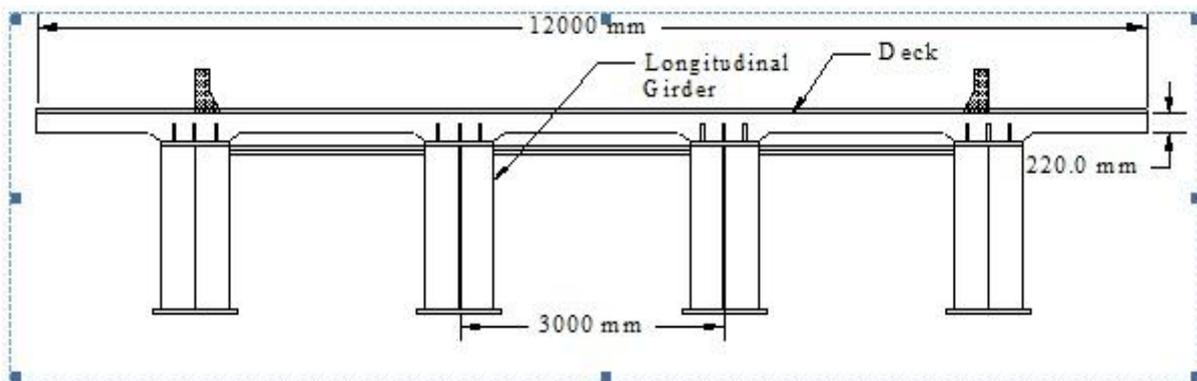


Fig. 3.b Cross Section of 4-Girders Composite Bridge

Various combinations of cross-section were generated to optimize the resulting bridge profiles keeping the maximum flexural stresses and maximum deflection at the mid span within the permissible limits. Resulting bridges were studied to investigate the influence of prestressing force and girder spacing.

III. DESIGN RESULTS

As per Indian Standards, Class 70R wheeled and tracked loads, two lanes Class A load, and Bogie load are considered for calculating the live load effects on the bridges. Maximum bending moment and deflection are computed using STAAD.Pro V8i software. Various results

for different spans and girder spacing are calculated using MATLAB.

Summaries of resulting designs are presented in Tables I and II, for 5-girders and 4-girders composite bridge designs, respectively.

It may be noted that these designs were performed to observe qualitative trends between the variables described above. Changes in the design assumptions will naturally change the resulting design values. Following subsections provide summary comments regarding the influence of the variable parameters on bridge performance.

Changing the tendon force and eccentricity alters the vertical force exerted on the structure. Vertical component of prestressing force is transferred to the support

through the vertical stiffener. Horizontal component of the prestressing force provides axial compression in the deck slab concrete.

Prestressing force for zero deflection under dead load and imposed load eliminate shrinkage cracks and increase the stress level in the deck slab, Further, higher stress level in the deck slab results in its better performance under fatigue loading.

Table-I. Summary of Study: 5 Girder System Composite Bridge

Span (m)		20	40	60	80	100
Girder web (mm)	Depth	1400	1800	2200	2600	3000
	Thickness	10	12	14	15	15
Girder top flange (mm)	Width	400	500	600	800	800
	Thickness	20	30	40	50	60
Girder bottom flange (mm)	Width	500	600	800	900	1000
	Thickness	20	30	40	50	60
Girder cross section area (mm ²)		32,000	54,600	86,800	124,000	153,000
Total Steel Area (5*Ag)		160,000	273,000	434,000	620,000	765,000
Without Hogging						
Total Prestressing force required in each cable (kN)		1080.5	2150.8	3423.2	4531.5	6000.3
Stresses in deck slab at mid section (N/mm ²)	Direct	3.27	6.31	6.86	6.26	7.19
	Flexural	5.84	7.33	7.48	7.20	7.04
	Total	3.27-9.12	6.31-13.64	6.86-14.35	6.26-13.54	7.19-14.22
Hogging with the 50% of LL deflection						
Total Prestressing force required in each cable (kN)		1322.3	2492.8	3819.5	4963.7	6457.9
Stresses in deck slab at mid section (N/mm ²)	Direct	5.31	8.24	8.38	7.55	8.31
	Flexural	5.84	7.33	7.48	7.20	7.03
	Total	2.39-8.24	4.57-11.90	4.64-12.12	3.95-11.15	4.79-11.83

Table-II. Summary of Study: 4 Girder System Composite Bridge

Span (m)		20	40	60	80	100
Girder web (mm)	Depth	1500	2000	2500	3000	3500
	Thickness	10	12	14	15	15
Girder top flange (mm)	Width	400	500	600	800	800
	Thickness	20	30	40	50	60
Girder bottom flange (mm)	Width	500	600	800	900	1000
	Thickness	20	30	40	50	60
Girder cross section area (mm ²)		33,000	57,000	91,000	130,000	160,500
Total Steel Area (4*Ag)		132,000	228,000	364,000	520,000	642,000
Without Hogging						
Total Prestressing force required in each cable (kN)		1170.4	2287.3	3549.3	4587.7	5959.4
Stresses in deck slab at mid section (N/mm ²)	Direct	2.90	5.47	5.71	4.91	5.55
	Flexural	5.96	7.20	7.23	6.91	6.72
	Total	2.90-8.87	5.47-12.67	5.71-12.94	4.91-11.82	5.55-12.27
Hogging with the 50% of LL deflection						
Total Prestressing force required in each cable (kN)		1441.8	2660.9	3975.4	5047.6	6442.6
Stresses in deck slab at mid section (N/mm ²)	Direct	4.87	7.29	7.13	6.11	6.60
	Flexural	5.96	7.20	7.23	6.91	6.72
	Total	2.01-7.85	3.69-10.89	3.51-10.74	2.65-9.56	3.24-9.96

3.1. Span Comparison for PSCC 5-Girder System Bridges

As the span length increases from 20.0 m to 100.0 m, the corresponding depth of girders required increases from 1400 mm to 3000 mm. The total cross-section area of girders required also increases from 160,000 mm² to 765,000 mm² with increase in the span length.

The prestressing force required in the cable increases from 1080.5 kN to 6000.3 kN for no hogging case, and from 1322.3 kN to 6457.9 kN for the 50% of LL hogging case. Thus, by marginally increasing prestressing force in the cables for the 50% of LL hogging case, the range of flexural stresses due to live load becomes half, which decrease the strain range in the concrete, so there will be very less fatigue in the deck slab.

Due to prestressing force the stresses in the deck slab are raised in the range of 3.27 N/mm² to 14.22 N/mm² for no hogging case, and 2.39 N/mm² to 11.83 N/mm² for the 50% of LL hogging case. The maximum stress in the hogging case is also lower than that of no hogging case. Higher stress level in the deck slab results in its better performance under fatigue loading. Shrinkage and creep, and prestressing losses are also taken care of by anchoring the cables into the deck slab.

3.2. Span Comparison for PSCC 4-Girder Bridges

As the span length increases from 20.0 m to 100.0 m, the corresponding depths of girders increase from 1500 mm to 3500 mm. With increase in the span length the total required cross-section area of girders increases from 132,000 mm² to 642,000 mm².

The prestressing force required in the cable increases from 1170.4 kN to 5959.4 kN for no hogging case, and from 1441.8 kN to 6442.6 kN for the 50% of LL hogging case. Thus, by marginally increasing the prestressing force in the cables for the 50% of LL hogging case, the range of flexural stresses due to live load becomes half, which decrease the strain range in the concrete resulting in lower fatigue in the deck slab.

Due to prestressing force the stresses in the deck slab are raised in the range of 2.90 N/mm² to 12.94 N/mm² for no hogging case, and 2.01 N/mm² to 10.89 N/mm² for the 50% of LL hogging case. The maximum stress in the 50% of LL hogging case is also lower than the no hogging case.

Higher stress level in the deck slab results in its better performance under fatigue loading. Shrinkage and creep, and prestressing losses are also taken care of by anchoring the cables into the deck slab.

3.3. Girder Spacing Comparison: 5-Girder System Vs 4-Girder System

For all the span length bridges the total area of steel required in the 4-girder system is nearly 20% lower than 5-girder system.

The required prestressing force in the cable is also marginally lower in the 4-girder system than 5-girder system.

The maximum stress (10.89 N/mm²) in the deck slab is lower in the 4-girder system in comparison to the 5-girder system (12.12 N/mm²).

IV. CONCLUSION

This study has presented the comparison of the total area of steel girder and prestressing force required in the cables, and stresses in the deck slab using various span lengths and girder spacings.

The following main conclusions are drawn from the study.

- (1) 4-girder system is found to be beneficial and economical in bridge design as compared to 5-girder system for all the span length bridges.
- (2) In comparison to no hogging case, in the 50% of LL hogging case, the range of flexural stresses due to live load is half, which decrease the strain range in the concrete, and hence results in reduced fatigue in the deck slab.
- (3) Shrinkage strain can well be taken care of by anchoring the tendons into end block of the deck slab. Further, by doing so stress level of concrete deck is raised, resulting in its better performance under fatigue loading.
- (4) In all cases, the 4-girder bridge case resulted in approximately 20% lower girder area (or weight) than the 5-girder bridge case.
- (5) In all cases, the prestressing force required in the 4-girder system bridge is little lower than that of 5-girder bridge system.
- (6) The maximum stress (10.89 N/mm²) in the deck slab is lower in the 4-girder system in comparison to the 5-girder system (12.12 N/mm²).

REFERENCES

1. Aravinthan T. (1999), "Flexural Behaviour and Design Methodology of Externally Prestressed Concrete Beams", *PhD thesis*, Saitama University, Saitama Japan.
2. Babaei, Khossrow and Fouladgar Amir, (1997), "Solutions to Concrete Bridge Deck Cracking", *Concrete International*, V. 19, No. 7, July.
3. Bentz D. P. and Jensen O. M., (2004) "Mitigation strategies for autogenous shrinkage cracking", *Cem. Concr. Compos.*, 26(6), 677-685.
4. Carlos Sousa1, Helder Sousa, Afonso Serra Neves and Joaquim Figueiras, (2012) "Numerical Evaluation of the Long-Term Behavior of Precast Continuous Bridge Decks", *Journal of Bridge Engineering*, Vol. 17, No. 1, 89-96.
5. Daly A. F. and Witarnawan W. (1997), "Strengthening of bridges using external post-tensioning", *Transport Research Laboratory*, Birkshire, U.K.
6. DD ENV-1992-1-1 Eurocode 2, "Design of Concrete Structures", British Standards Institute.
7. Giussani F., Mola F. and Palermo A., (2003) "Continuity effects in composite steel-concrete railway bridges" *Proc. 6th Int. Conf. Maintenance and Renewal of Permanent Way, Power and Signaling; Structures and Earthworks*, Engineering Technics Press, Edinburgh, UK.
8. IRC: 22-1986, "Code of Practice for Road Bridges, Section VI, Composite Construction." Indian Road Congress, New Delhi.

9. IRC: 6-2000, "Standard specifications and Code of Practice for Road Bridges, Section II, Loads and Stresses", Indian Road Congress, New Delhi.
10. IS 1343:1999, "Code of Practice for Prestressed Concrete", Bureau of Indian Standards, New Delhi.
11. IS 2062: 1999, "Steel for General Purposes" Bureau of Indian Standards, New Delhi.
12. IS 800: 2007, "General Construction in Steel - Code of Practice" Bureau of Indian Standards, New Delhi.
13. Kasim.S.Y and Chen.A. (2006), "Conceptual design and analysis of steel concrete composite bridges" *Technical article*.
14. Kim H. Y., Jeong Y. J., Kim J. H. and Park S. K. (2005), "Steel concrete composite deck for PSC Girder bridges." *Journal of Civil Engineering, ASCE*, Vol-9, No.-5, September, pp.385-390.
15. Miyamoto A., Tei K., Nakamura H., and Bull J. W. (2000). "Behavior of Prestressed Beam Strengthened with External Tendons." *Journal of Structural Engineering*, 126(9), 1033-1044.
16. Newhouse C. D., Roberts-Wollmann C. L., Cousins T. E., and Davis, R. T., (2008) "Modeling early-age bridge restraint moments: Creep, shrinkage, and temperature effects" *J. Bridge Eng.*, 13(5), 431-438.
17. Rambod Hadidi and M. Ala Saadeghvaziri, (2005) "Transverse Cracking of Concrete Bridge Decks: State-of-the-Art", *Journal of Bridge Engineering*, Vol. 10, No. 5, 503-510.
18. Ryu H. K., Kim Y. J., and Chang S. P., (2007) "Crack control of continuous composite two girder bridge with prefabricated slabs under static and fatigue loads", *Eng. Struct.*, 29(6), 851-864.
19. Sandeep Chaudhary, Umesh Pendharkar and Ashok Kumar Nagpal, (2009) "Control of Creep and Shrinkage Effects in Steel Concrete Composite Bridges with Precast Deck", *Journal of Bridge Engineering*, Vol. 14, No. 5. 336-345.
20. Singh P. K., (2008), "Fatigue in Concrete Decks of Cable Stayed Bridges", *Proc. Int. Conf. on 'Innovations in Structural Engineering and Construction'*, Taylor-Francis Group, London.

Fe Analysis of Reactive Powder Concrete in Direct Shear

Mr. M K Maroliya.

Assistant professor, Applied Mechanics Dept, Faculty of Technology & Engineering, M.S. University of Baroda, Vadodara,

ABSTRACT : Steel fibre reinforced concrete is a new building material although the cradle of intense research traces back to the seventies. In order to investigate the possibilities of current FE programs for implementation of steel fibre concrete this study was conducted. The aim of this study was to show the use of steel fibre concrete with FE calculations, to point out the possibilities and restrictions and to give recommendations for practical usage. Because of the complexities associated with the development of rational analytical procedures, present day methods continue in many respect to be based on the empirical approaches using result from a large amount of experimental data. The finite element (FE) method offers a powerful and generic tool for studying the behaviour of structures. The power of the FE method is in its versatility. The structure analyzed may have arbitrary shape, arbitrary supports and applied tractions. Such generality does not exist in classical analytical methods, even when the structural geometry is simple.

Key words: RPC, non linear analysis, ATENA, shear

I. INTRODUCTION

Today's complex construction being planned by a civil engineer is predominantly calculated by means of Finite Element (FE) programs. For this reason it is quite evident that new building materials have to be considered in FE programs to support fast acceptance of them in practice. It must be emphasized that the programs are rarely used by a civil engineer in practice. This is mainly due to the fact that they are often too complex for general applications that are the time for getting to know the work mode does not justify the advantages you get. Furthermore most applications demand merely for linear calculations that is material non-linearity is not regarded.

The behavior of members and structures, specifically their response to loads and other actions, has been the subject of intensive investigation for many years. However, challenges in designing complex structures has prompted the structural analyst to acquire a sound understanding of the structural behaviour of concrete structures. In many cases, in non-conventional designs, code provisions cannot be relied upon to provide realistic information on design issues such as the load-displacement response and strength and failure modes of a structure and/or its structural elements. Such information is required for safe and cost effective design. In concrete and reinforced concrete structures, the behaviour under loads can be complex.

More specifically for concrete issues such as confinement, cracking, tension stiffening, non-linear multi-axial material properties, complex steel-concrete interface behaviour and effects, that are commonly treated in an approximate way, the FE method has thus become a powerful tool that allows the analyses of complex structures and structural phenomena.

While simplified 2D analysis can fully represent all the aspects of the response of structures and structural elements where triaxial behaviour is important. With rapid developments in numerical techniques, non linear 3D analysis of structural elements has become increasingly available.

II. MODELING

Modeling is the primary task of any analytical study and the result obtained to a large extent depends on the simplification taken during in this step. Modeling involves creation of geometry of overall structure by including elements of various components representing respective structural behavior including boundary conditions in the considered problem, Material properties of various elements and loads on various elements and its combinations were defined.

III. NON-LINEAR ANALYSIS

Once the model is built, the non-analysis is performed after defining the various loads and supports. The CC3NONLINEAR CEMENTITIOUS2 is suitable for fibre reinforced concrete, such as SHCC (strain hardening cementitious composites). Tensile softening regime and the shear retention factor are modified based on the model, proposed in KEBELE, P. This model is based on a notion of a representative volume element (RVE) which contain distributed multiple crack (hardening) as well as localized cracks (softening) the overall strain of the RVE is then obtained as a sum of strain of material between cracks (which may possibly contain non linear plastic strain due to compressive yielding), cracking strains due to multiple cracks, and cracking due to localized cracks. In this study an inverted 'L' shape specimen with varying depth of shear plane is studied and results are compared with the experimental results by keeping same loading and material of RPC.

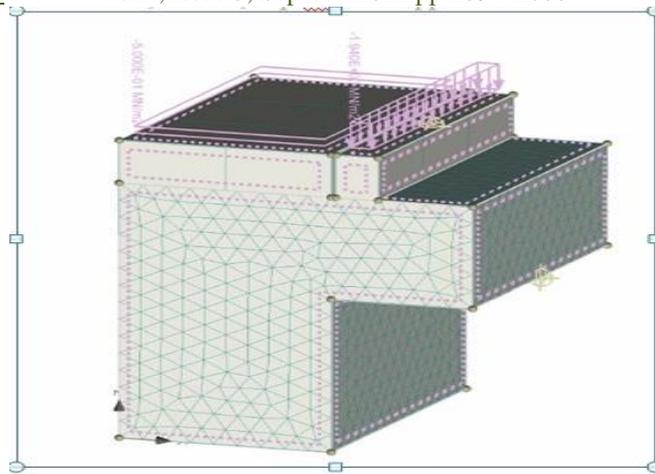


Fig.1.0 General Arrangement of specimen with FE mesh.

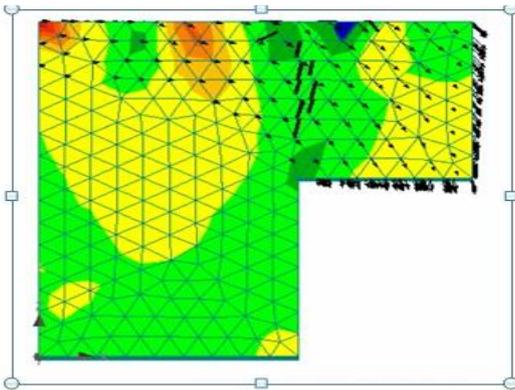


Fig.2.0 Crack Propagation In Specimen.

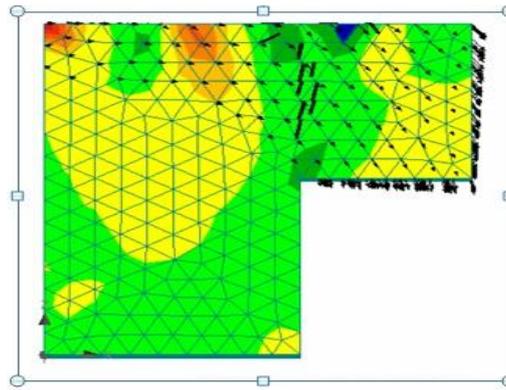


Fig. 3.0, Displacement in a Specimen.

Table .1.0 material properties used in ATENA.

MATERIAL-1	
Title	3D Nonlinear Cementitious 2
Type	CC3DNonLinCementitious 2
Elastic Modulus, E (MPa)	4.452E+4
Poisson's ration, μ [-]	0.200
Specific material weight, ρ [MN/m ³]	2.300E-02
Coefficient of Thermal Expansion, α [1/K]	1.200E-05
Tensile Strength, F_t [MPa]	7.071E+00
Compressive Strength, F_c [MPa]	-1.360E+02
Specific Fracture Energy, G_f [MN/m]	1.694E-04
Critical Compressive Displacement, W_d [m]	-5.000E-04

Exc., def. the shape of fail. surface, e [-]	0.520
Multiplier for the direction of the pl. flow, β [-]	0.000
Fixed crack model coefficient [-]	1.000
Plastic strain at compressive strength, ϵ_{cp} [-]	-2.833E-03
Onset of non-linear behavior in compression F_{co} [MPa]	-1.485E+01
Reduction of compressive strength due to cracks, c [-]	0.8
MATERIAL-2	
Title	3D Elastic Isotropic
Type	CC3DElastIsotropic
Elastic Modulus, E (MPa)	2.100E+05
Poisson's ration, μ [-]	0.300
Specific material weight, ρ [MN/m ³]	2.300E-02
Coefficient of Thermal Expansion, α [1/K]	1.200E-05

IV. RESULTS AND DISCUSSION

To the presence of dominant shear stresses, the specimen geometry was studied using finite element analysis, the stress concentration are obtained as shown in **Fig.5.4** indicating the possibility of cracking. The shear stress distribution is shown in **Fig. 5.4** with darker zone indicating higher stresses. It can be seen as that there is a zone of shear stresses at the junction of "L" shape specimen, conforming that shear failure would occur if tensile cracking was to be restrained by fibres. As a result clear shear crack propagation can be observed at the junction of "L" shape specimen (**Refer Fig.5.2**).

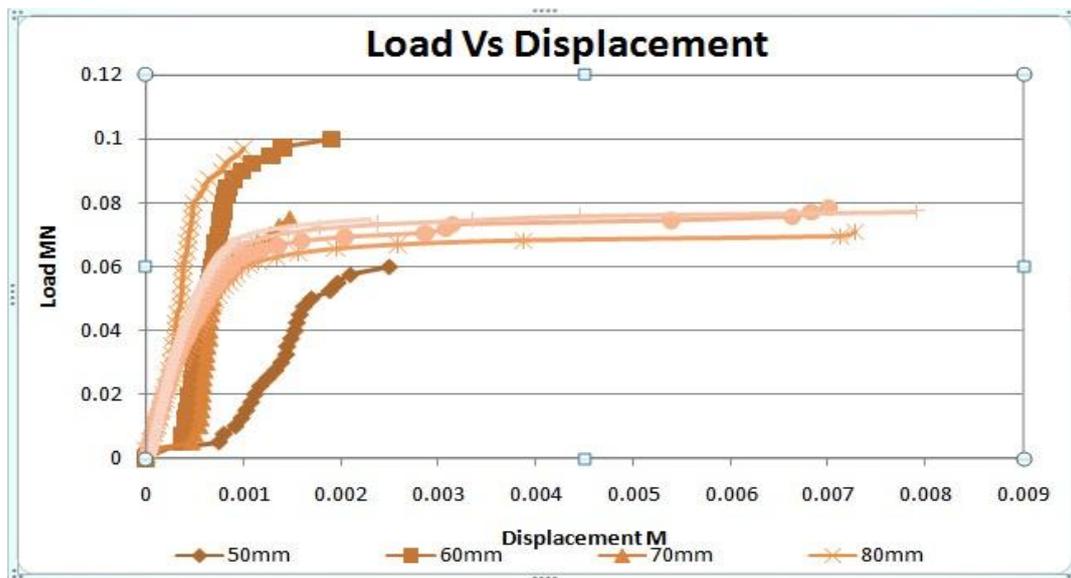


Fig. 4.0, Superimposed Load Displacement Relation with Thickness of Shear Plane Using ATENA and experimental values. .

The load deformation results of the FE modeling for 60 mm and 70 mm thickness of shear plane generally compare well with the test data. The FE results of the 80 mm grossly over predict the failure load. It is worth mentioning that 50 mm

thickness of shear plane fail somewhat prematurely because of the localized crushing and tensile failure. A good co-relation between results of FE analysis experimental values can be obtained using "ATENA" having powerful simulation capability. It is using model propose by KEBELE, P. the model is based on notion of representative volume element which contain distributed multiple crack as well as localized crack.

Looking at the displacement corresponding to peak load, it would be expected that the influence of the existence of fibres on the response of the RPC would be to induce some degree of plastic behavior.

The toughening observed in the test is a result of the pull out and dowel action of the fibre during the shear cracking. The pull out resistance tends to close the cracks, while the shear forces tends to open it due to the irregularity of the crack face as illustrated in **Fig.3.0** this produces a confining effect which together with the dowel action of the fibres and friction between the crack faces, can lead to an increase in the shear strength.

V. CONCLUSIONS

FE analysis is done using ATENA, ATENA is having inbuilt material model for HPFRC which can simulate behaviour of RPC also ,the results obtained by FE analysis are well compare with the experimental results .

Optimum thickness of shear plane to be considered is 60 mm which is getting conformed by FE analysis. Initially RPC shows linearly elastic behaviour and then dominated by plastic behaviour representing the behaviour of Elasto-plastic material.

REFERENCES

- 1) *Concrete Membrane Elements in Shear*". ACI Structural Journal, Vol.92, No. 6, pp. 1-15.
- 2) Yen lei Voo, Stephen j. Foster and R. Ian Gilbert (2006) "*Shear Strength Of Fiber Reinforced Reactive Powder Concrete Prestressed Girders Without Stirrups*". Journal of Advanced Concrete Technology Vol. 4, No. 1, pp. 123-132.
- 3) Russel, H G (1999) "*ACI defines High -performance concrete*", Concrete International, Vol. 21, pp. 56-57
- 4) Richard Pierre, Marcel Cheyrezy (1995) "*Composition of Reactive Powder Concretes*". Cement and Concrete Research, Vol. 25. no. 7, pp. 1501-1511
- 5) Oliver Bonneau, Claude Poulin, Jerome Dugat, P, Richard, and Pierre-Claude Aitcin (1999) "*Reactive Powder Concretes: From Theory to Practice*". Concrete International. pp.47-49.
- 6) Khaloo Ali R. and Nakeseok kim. (1997) "*Influence Of Concrete And Fibre Characteristics On Behavior Of Steel Fibre Reinforced Concrete Under Direct Shear*". ACI Materials Journal, Vol. 94, No. 6, pp. 592-664.
- 7) Kazunori Fujikake, Takanori Senga, Nobuhito Ueda, Tomonori Ohno and Makoto Katagiri. (2006) "*Nonlinear Analysis For Reactive Powder Concrete Beam Under Rapid Flexural Loadings*". Journal of Advanced Concrete Technology Vol. 4, No. 1, pp. 85-97.
- 8) Francesco Bencardino; Lidia Rizzuti; Giuseppe Spadea; and Ramnath N. Swamy (2008) "*Stress-Strain Behavior of Steel Fiber-Reinforced Concrete in Compression*" Journal of Materials in Civil Engg. ASCE, Vol 20 No. [3] pp. 255-263.
- 9) Chau K.T., X.X. Wei (2007) "*Finite Solid Circular Cylinders Subjected To Arbitrary Surface Load. Part I – Analytic Solution*". International Journal of Solids and Structures 37, pp. 5707-5732.
- 10) Chau K.T., X.X. Wei (2007) "*Finite Solid Circular Cylinders Subjected To Arbitrary Surface Load. Part I – Analytic Solution*". International Journal of Solids and Structures 37, pp. 5707-5732.
- 11) Bairagi N. K. and Modhera C. D. (2001) "*Shear Strength Of Fibre Reinforced Concrete*". ICI Journal, Vol. 1, No. 4, pp. 47-52.
- 12) Abouzar Sadrekarimi. (2004) "*Development of a Light Weight Reactive Powder Concrete*". Journal of Advanced Concrete Technology Vol. 2, no. 3, pp. 409-417.
- 13) Byung-Wan Jo, Chang-Hyun Kim, (2007) "*Characteristics Of Cement Mortar With Nano-Sio2 Particles*", construction and building Materials 21 pp. 1351-1355.
- 14) Chakraborti A.K., I. Ray, B.Sengupta (2001) "*High Performance Concrete for Containment Structures*". Transactions SMiRT 16, Washington DC, August 2001, pp.1-8.
- 15) Dias, WPS, *et al*, (1990). "*Mechanical Proprieties Of Hardened Cement Paste Exposed To Temperature Up To 700°C (1292°F)*". ACI Material Journal, 87(2): pp. 60 - 165.

Experimental Study on Al₂O₃-R134a Nano Refrigerant in Refrigeration System

D. Sendil Kumar¹, Dr. R. Elansezhian²

¹Research Scholar, Department of Mechanical Engineering, Pondicherry Engineering College, Puducherry, India

²Associate Professor, Department of Mechanical Engineering, Pondicherry Engineering College, Puducherry, India

ABSTRACT : In this paper, an experimental work was investigated on nanorefrigerant. Nano Al₂O₃-PAG oil was used as nano-refrigerant in R134a vapour compression refrigeration system. An experimental setup was designed and fabricated in the lab. The system performance was investigated using energy consumption test and freeze capacity test. The results indicate that Al₂O₃ nano refrigerant works normally and safely in the refrigeration system. The refrigeration system performance was better than pure lubricant with R134a working fluid with 10.32% less energy used with 0.2%V of the concentration used. The results indicate that heat transfer coefficient increases with the usage of nano Al₂O₃. Thus using Al₂O₃ nanorefrigerant in refrigeration system is found to be feasible.

Keywords: Nanorefrigerant, Al₂O₃ nanoparticle, COP, performance

I. INTRODUCTION

HFC134a is the most widely used alternative refrigerant in refrigeration equipment such as domestic refrigerators, chillers and automobile air conditioners. HFC134a has been accepted as long term alternative refrigerant in many countries. Due to strong chemical polarity, the normally used mineral oil is replaced by Polyol-ester (POE) oil. Nano-particles suspended in lubricant of the compressor is found to be increasing the system efficiency and also the system works without any choking. Nano Al₂O₃ gives better performance by returning more lubricant oil back to the compressor. Eed et al. investigated the evaporating heat transfer co-efficient and heat flux in a domestic refrigerator. The presence of nano CuO in the refrigerant R1234a improved the evaporating heat transfer coefficient with the particle size from 15 to 70 nm. Sheng et al. studied the reliability and performance of a domestic refrigerator with TiO₂ and mineral oil lubricant with R134a. The results showed reduction in 26.1% energy consumption. Ching-Song Jwo et al. studied the usage of alumina in R134a refrigeration system. The thermal conductivity was found to increase from 1.5 to 4.6%, when the sample temperatures were 20°C to 40°C at 1.5 wt.%. Park investigated the effect of CNTs on the nucleate boiling heat transfer of R123 refrigerant.

II. EXPERIMENTAL SETUP

2.1 COMPONENTS

The experimental consists of compressor, fan cooled condenser, expansion device and an evaporator section. Capillary tube is used as an expansion device. The

Evaporator is of coil type which is loaded with water. Service ports are provided at the inlet of expansion device and compressor for charging the refrigerant. The mass flow rate is measured with the help of flow meter fitted in the line between expansion device and drier unit. The experimental setup was placed on a platform in a constant room temperature. The ambient temperature was ±1.5°C. The air flow velocity was found to be less than 0.35m/s. Table1. shows the specification of the refrigeration system

Table 1. Refrigeration system specifications

Refrigerant	R134a
Charged mass	150gm
Compressor type	Reciprocating
Gross capacity	125 L

2.2 Instrumentation

The temperatures at different parts of the experimental setup are measured using resistance thermocouples. 12 resistance thermocouples were used for the experimentation. The pressure at compressor suction, discharge, condenser outlet and at evaporator outlet are measured with the help of pressure gauges. The power consumption of the system was measured by a digital Watt-hr meter. A digital wattmeter is also connected with the experimental setup. Table2. summarized the characteristics of the instrumentation.

Table 2. Measurement Equipment

Variable	Device	Range
Temperature	Pt100 PID controller	-50 to 199°C
Pressure	Pressure Gauge	0-10 bar
Power	Digital Watt/Watt-h meter	5-20A

III. EXPERIMENTAL PROCEDURE

3.1 Preparation of nano- Refrigerant

Nanoparticles of Al₂O₃ are added to the refrigeration system by adding them to the lubricant in the compressor of the system. The preparation and stability of this

lubricant and nanoparticle mixture is very important. The lubricant oil, a type commonly used in refrigeration and air-conditioning systems was poly alkylene glycol (PAG). This oil is selected owing to its common usage and superior quality.

The nanoparticles of Al_2O_3 in the range 40-50 nm were mixed with PAG to synthesize nanolubricant in a recommended method for nanofluid. PAG oil was used as supplied by supplied without further purification. The nano particles of Al_2O_3 and PAG mixture was prepared with the aid of magnetic stirrer for 2 hrs. The mixture is then further kept vibrated with an ultrasonic homogenizer for half an hour to fully separate the nanoparticles and to prevent any clustering of particles in the mixture to obtain proper homogenization. No surfactant is added in this work as there may be any influence in reduction of thermal conductivity and performance.

3.2 Nano- Refrigerant Concentration

Nano-Refrigerant with 0.2% concentration of Al_2O_3 in the refrigerant R134a is prepared and tested in the setup.

3.3 Charging of the set up

The fabricated experimental setup was filled with N_2 gas at a pressure of 5 bar to 7 bar and this pressure is maintained for 45 minutes. Thus the system was ensured for no leakages. The system was evacuated by removing N_2 gas. A vacuum pump was connected to the port provided in the compressor and the system was completely evacuated for the removal of any impurities. This process was carried out for all the trials. Through the service ports nano refrigerant was carefully added to the system. Precision electronic balance with accuracy $\pm 1\%$ was used to charge a mass of 150 gm. into the system. Every time the system was allowed to stabilize for 15 min.

IV. Performance Test

A performance test is made for 150 gm. of pure R134a system which is treated as the basis for comparison with other results. Nano Al_2O_3 -R134a with 0.2% concentration was fed to the experimental setup and the tests were conducted under the same conditions. In order to obtain repeatability each test was run for 3 to 4 times. Performance tests was also conducted with charge mass of the order of 150gm., 180 gm. and 200gm.

4.1 Factors affecting Refrigeration System

The important factors affect the performance of refrigeration system are Refrigeration effect, Coefficient of Performance (COP) and Energy factor(EF).

- a) Refrigeration effect
 $q = \text{Heat removal} / \text{mass flow rate or refrigerant}$ (1)
- b) Coefficient of Performance
 $\text{COP} = \text{Heat Removal} / \text{Work Input}$ (2)
- c) Energy Factor
 $\text{EF} = \text{Cooling capacity} / \text{Power consumption}$ (3)

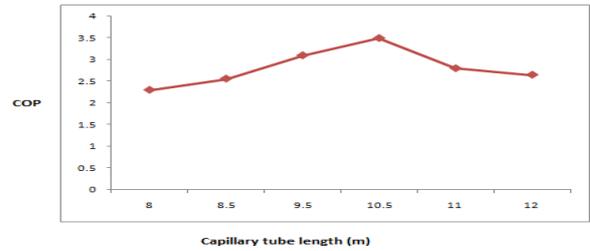


Figure 1. Variation of COP with capillary tube length

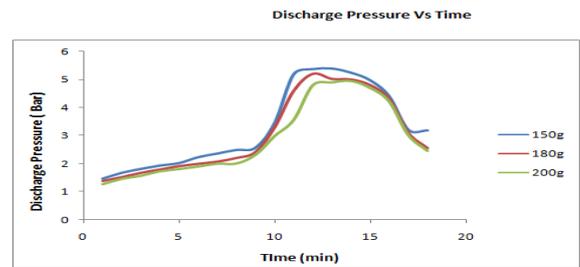


Figure2. Variation of Discharge pressure with time

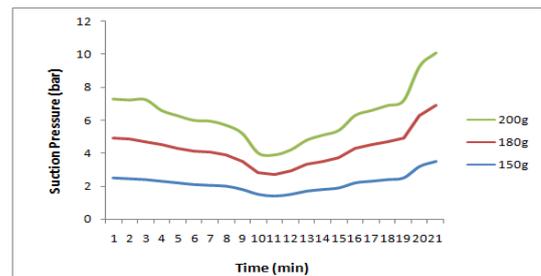


Figure 3. Variation of Suction pressure with time

V. CONCLUSION

Addition of Nano Al_2O_3 in to the refrigerant shows improvement in the COP of the refrigeration system. Usage for Nano refrigerant reduces the length of capillary tube and cost effective. The following are the important results of the experimental investigation.

1. The COP of the system increases with increase in capillary tube length
2. Maximum COP of 3.5 is achieved for a capillary length of 10.5 m
3. The Discharge pressure increases with time and attains a maximum value and then decreases.
4. The Maximum discharge pressure is obtained for charge mass of 150gm.
5. The suction pressure decreases initially and then increases with time.
6. Suction pressure is found to be less for a charge mass of 150gm.

References

- [1] Bolaji, B.O. (2010). "Experimental Study of R152a and R32 to replace R134a in a domestic refrigerator," *Int J of Energy*, Vol 35, pp 3793-3798.

- [2] Mohanraj, M , Experimental investigation of R290/R600a mixture as an alternative to R134a in a domestic refrigerator, *Int J of Thermal Sciences*, Vol. 48,2009, 1036-1042.
- [3] Ching-Song Jwo, Experimental Study on Thermal Conductivity of Lubricant containing Nanoparticles," *Rev Adv Mater Science*, Vol. 18, 2008 ,660-666.
- [4] Avinash, Development of Energy efficient R134a Automotive Air-conditioning system using Internal Heat Exchanger, *Proceedings of World Congress on Engineering*, Vol. III., 2011
- [5] Kai-Shing Yang, An investigation of a Top-mounted domestic refrigerator, *Int J of Energy Conservation and Management*, Vol. 51, 2010, 1422-1427.
- [6] Shengshan Bi, Performance of a Domestic Refrigerator using TiO₂-R600a nano-refrigerant as working fluid, *Int J of Energy Conservation and Management*, Vol. 52, 2011, 733-737
- [7] Saidur, R, A Review on the performance of nano-particles suspended with refrigerants and lubricating oils in refrigeration system, *Int J of Renewable and Sustainable Energy Reviews*, Vol. 15, 2011, 310-323.
- [8] Joseph Sekher, S. , Improved energy efficiency for CFC domestic refrigerators retrofitted with ozone friendly HFC134a, *Int J of Thermal Sciences*, Vol. 43, 2004 ,307-314.
- [9] Ching-Song Jwo, Efficiency analysis of home refrigerators by replacing hydrocarbon refrigerators, *Int J of Measurement*, Vol. 42, 2010, 697-701.

Quality Benefit Analysis of Software Automation Test Protocol

S Janardhana Rao¹, S. Ajay Kumar² Dr. R. Satya Prasad³, K. Eswara Rao⁴

¹(Team Leader, A P society for knowledge networks, Hyderabad, Andhra Pradesh, India)

²(S/w developer, Techmahindra, Chennai, India,)

³(Associate professor, Department of CSE, Acharya nagarjuna university, Andhra Pradesh, India)

⁴(Assistant professor, Department of CSE, AITAM, Tekkali, Andhra Pradesh, India)

ABSTRACT: *Software testing is very important to provide software quality and to improve the reliability. As the volume and complexity of software applications tend to increase, user's requirements for more quality software is also rising. The importance of software testing in the software development process is also increasing exponentially, while software testing effort also appears to be increasingly difficult. Hence software testing is still an open research area and effective software testing method is required to improve software quality. In this paper we analyze software automation testing, definition, characteristics and functions are discussed. We propose a method to improve the overall process of software automation testing.*

Keywords: *Software testing, Automation Protocol, software quality, software reliability*

I. Introduction

Software testing is an important stage in software life cycle. Quality testing directly determines the stable operation of software products. It is well known that software testing consumes more time with high labor intensity and easily-introduced artificial error therefore automation software testing method is very important. Automated testing technology in software testing is to be further improved, as work load of software testing is very large (accounting for about 40%-50% of overall development cycle), of which most of the work applies to automation, so the test improvement will bring very significant results to cost, quality and cycle of the whole software projects development work. Generally speaking, it is through the development of the automation tools and the execution of the testing scripts for automation testing to achieve the purpose of software quality evaluation.

In the software development process, first of all it requires that we verify by testing the software developed meet the demand analysis description in function, meet customer requirements of carrying load and the corresponding response time, computation requirements in performance; on the other hand, staff are also required to release software as soon as possible within the budget to reduce spending. Traditional software testing process usually carries out a small number of unit's tests first in the software development process, and then at the end stage of

the software development, focuses on a lot of testing, including functional and performance integration testing and system testing. As more and more complex of software projects, the software testing process brings the following issues for project developers

A. *The project process is difficult to control*

Project management becomes more difficult; a lot of software bugs are often not discovered until later stage of software testing, it is difficult to estimate the cost of expensive in solving the problem, so the project process cannot be controlled. In addition, in the software design process, relevant staff can not accurately understand the software quality state, potentially increasing the difficulty of the project management and development.

B. *Difficult to control project risk*

Project risks cannot be reduced until the development later stage. Usually after testing personnel carry out the system test, it is possible to determine whether the entire development process meet the system functionality, performance and reliability requirements of users.

C. *Project development costs exceed budget*

Throughout the project development cycle, the later errors are found, the higher repair costs, then the problem is not resolved in a timely manner will result in a rapid increase in the project cost.

The rest of the paper is organized as section 2: discuss about analysis of software automation test, section 3: presents improvements to software automation testing process, section 4: concludes the paper.

II. Analysis Of Software Automation Test

A. *The definition of software automation test*

Currently, the conventional definition of software automation test is "the use of an automated testing tool to verify the needs of a variety of software testing, which includes the management and implementation of testing activities." Usually, we carried out automated testing refers to an automated testing has used the formal process of manual testing process.

B. *Why should the introduction of automated software testing*

Typically, software testing is accompanied with many repetitive, non-intellectual and non-creative operations, computer can replace manual to accomplish this task, so that programmers are released from the complex and repetitive work and turn to a deeper level of project testing. Software automation testing is based on the existence of manual testing, mainly through the corresponding software testing tools, scripts, etc. to implement, has good maneuverability, repeatability and high efficiency characteristics. Automated testing is of high testing efficiency, low cost, it can automatically convert test with strong repeatability from manual to computer; very suitable for fast regression testing, change the situation that manual testing is difficult to make rapid assessment to the new iterative version; avoid common errors in manual testing, and test time-consuming boundary value can rapidly and accurately implement the concurrent operation of the multi-user that manual testing is difficult to achieve, automatic test is easy to complete. Therefore, the implementation of test automation is a development direction software testing industry. Mastering advanced technology in this field has a huge boost for a software company's core competitiveness and future development.

C. Common tools in software automation testing

There are many tools for software automation testing, some more commonly used tools are introduced in details given by the following the table below table 1.

Table 1: Tools in software automation testing

Unit testing	Java testing tool: Junit
Functional testing	Mercury Interactive Company WinRunner, QuickTest Pro, IBM Rational Company's Rational Robot, Rational Xde Tester, Compuware Company's QARun, RadView Company's WebFT, Empirix Company's eTester, etc.
Stress testing	Mercury Interactive Company's Winload, Compuware Company's QALoad
Load testing	Mercury Interactive Company's LoadRunner
Web testing tool	Empirix Company's eTest Suite
Web systematic testing tool	Microsoft Company's Web Application Stress Tool
Regression testing	Rational Company's Team Test
Performance testing	Mercury Interactive Company's Loadrunner, Compuware Company's QALoad, Microsoft Company's web Application Stress Tool, etc.
Testing process management tool	IBM Rational Company's Rational Testmanager, Mercury Interactive Company's TestDirector
Defect tracking tool	Compuware Company's senior changes and defect management tool: TrackRecord

White box testing	IBM Rational Company's Rational Purify pluS, Compuware Company's DevPartner Studio Professional Edition
Black box testing	QACenter, SQATeamtest, Rational Visual Test

D. Design test scripts

Script is a set of code that is executed by test tools, it can be generated by recording the operation of the test, and then do modifications, thus reducing the workload of scripting. Of course, it can directly use scripting language to write scripts. Scripting technologies can be divided into linear script, structured script, sharing scripts, data-driven scripts and keyword driven script. The principles of editing script is try to cover most test items, with strong compatibility, making a new project testing can be easily inherited and modify the original script, to reduce the subsequent workload.

III. Improvements to Software Automated Testing Process

A. Determine the feasibility of products automated testing

Automated testing can improve test efficiency, but for short period, tight scheduled projects, automated testing should not be used. Preparatory work for the implementation of automated testing is very large, to apply enterprise-level automated testing framework into a project needs to assess its suitability, it must not be blindly applied to a test project, in particular, not suitable for project with short cycle, because it is very likely to need a lot of test preparation and implementation of the framework, which will lead to insolent. For example, for a unit test automation, and develop hundreds of lines of test a code line with only 10 rows of function, regardless the maintenance costs of developed automated testing, at least you can have an intuitive understanding: shift from manual testing to automated testing has costs, and this investment is far higher than cost of manual testing, it must firstly to determine whether automated testing of a software project is appropriate.

B. Make risk assessment for software project testing

Financial risk, although some projects have automated testing conditions, organizational restructuring cost estimates after the introduction of automated testing is necessary. Entry mode risks must be considered with automated testing and manual testing combination, design appropriate ratio to ensure smooth testing. At the start, 80% tests can be set to be manual testing, when these goals are achieved, then increase the automated testing rates. Time estimates, after finishing the assessment of above several indicators, it is necessary to estimate the time period to

implement test automation in order to prevent unnecessary waste of time, to reduce the unwarranted consumption in personnel, capital and resources. Although after the test automation goes in the right track, it will play a multiplier effect, the early investment is huge, it is required to fully consider all the factors, clarify implementation plan and strictly according to plan, in order to minimize risk.

Workflow change risk. The test team or even the whole development organization implement test automation, more or less due to adapt to the work process of test tools, bringing the corresponding changes in team's test process, development process. And, if the change is unreasonable, it will cause much complain mood of team members, so it should minimize the changes, and to overcome the difficulties that may exist in changes.

C. Develop careful test plan

Test plan includes the establishment of test environment, determining the test strategy, testing design process control and fault-tolerant processing. According to characteristics of developed software project, develop careful plan, design various steps of test. Due to automated test has relatively high requirements on software design quality, it cannot do anything to emergencies and issues cannot be solved by the software, so solutions should be determined in advance for test abnormality, the problem can be resolved quickly, reduce unnecessary overhead. Automation test strategy refers to make sure which tests execute automated testing and when to adopt the automation testing. One of the misunderstandings for people is that they should realize the probability of the automated testing to the most degree, that is, the earlier the better; usually people believe that the higher the automated testing is, the higher the utilization rate of the automated testing instrument is and the greater we can obtain the investment rewards. In fact, the profits we obtain from test automation tool should be reflected in the test on quality, not quantity, and what tests should be chosen to implement the automation and how to develop executive test scripts are more important than how many tests has been automated. How do people select the best way for the execution of automated test? Usually it can be divided into three steps, that is, the extraction of testing suitable for automation, the evaluation of the time consumption for each automation testing and the verification of the order for the automation testing priorities according to the test goal. Firstly, make the tablet to extract automation testing project with the principle of selecting the testing item which can get the biggest investment rewards, namely, to shorten the time period most, to reduce the risk most and to improve the test accuracy most. Then, evaluate the time consumption for the automation test. Presently there is not a simple mathematical model to test the proportion between the time consumptions for the automation testing and manual testing; however, according to the estimation of the test expert, the time of an

automation testing for software is thrice to ten times as much as a manual testing, for complex testing, or even longer. Therefore, if a test suite need 100 hours to test, for automation test it will takes 300 to 1000 hours or more time, and certainly we have to admit that any estimation is only a surmise. Although we should make the judgment on the basis of enterprise testers' actual testing skills, test software actual characteristics and the actual testing tool-use complexity, one thing is uncountable that the time consumption by the first implementation automated test is longer than the time by the familiarity with tools and the testing, so in evaluation of the time consumption by automation testing we must take it into consideration. My basic principle is to select the test with the longer time consumption to execute automation test firstly.

D. The combination of two programs: manual testing and automated testing

Test automation can bring very significant benefits, but we cannot completely rely on automated testing, it is only a part of testing, is a complement to manual testing. Automated testing is no substitute for manual testing; both have their own characteristics, with different test object and test range. According to report, automated testing can only detect 15% of defects, and manual testing can detect 85%. Before automated testing, we should first establish a concept of software automation testing. Unit testing, integration testing, system load testing, performance testing, stability testing, reliability testing are suitable for automated testing. Instable software testing, software with short development cycle, one-time software is not suitable for automated testing. We can choose according to actual situation. In most cases, manual testing and automated testing should be combined, to complete the testing tasks in the most effective way. During functional testing, the precise meaning of automated testing tool is regression testing tool, at that time the tool cannot find more new problems, but can guarantee accuracy and objectivity of parts that have been tested

E. Summarizes the process of software automated testing

To record the entire testing process, summarize experiences and lessons learned in the process, generate test reports, find ideas and methods for resolving similar problems, so in the future testing similar software projects can reuse the previous results achieved, and if able to do so, on the basis of the test, design related test template, if the future has software projects with similar structure, we can follow this model to test, thereby greatly improve the efficiency of the test.

IV. Conclusion

In short, domestic development of automation test in software project is not very mature, has not really formed a whole set of reasonable and effective norms. In practice, we have to combine features of traditional manual testing

and automated testing, play their strengths, so that automated testing strategies and tools work as a weapon in the hands of testers, release them from boring and repetitive work, focusing time and energy on complex work requiring intelligent judgment and other new test cases. In addition, we must avoid equate automated testing and testers together, not asking too much for automated testing. We shall set a proper automated testing perception, have a clear understanding that manual testing is a powerful complement to automated testing and not replace the status of the tester. Any single technology or operation advances cannot independently ensure a large scale of improvements to software development efficiency, stability and maintainability in a short time. Automated testing is also an accumulation of experience, gradual process, software testers cannot be expected to automate all the tests in a short time. Successful automated testing requires developing appropriate automated test program, and reasonable automated testing is the first step in the success implementation of the automated testing strategy. Only by fully taking into account the risk of implementation of automated testing, resources and objectives, it is possible to develop automated testing strategies for your own, and

ultimately improve test efficiency and reduce cost of test. Each enterprises with testing department or engaged in testing business, should learn overseas and domestic advanced testing experience, refer to popular industry standard, find their own team's testing methods and models to create greater social value.

References

- [1] Yang Xina, Yuan Congling. "The Java-based Web Application Performance Testing", 2006.
- [2] Chen Hua. "The Accuracy of Testing Process", 2005.
- [3] Chen Yanzhuo. Metric Testing Tools - Three Indicators for Assessment of Automated Performance [4] Testing Tools. software world, 2005.
- [5] Software Automation Testing Technology. IBM Rational Technical White Paper.
- [6] Liu Sheng. Design and Practice of Software Automation Testing Framework. Posts & Telecom Press.
- [7] Li Yongzhong. Analysis of Effective use of Automated Testing in Software testing, 2006.
- [8] Richard Bender. Requirement-based Test is the Essence of SoftwareTesting, 2010.

Size Effect on Shear Behavior of Reactive Powder Concrete Containing Steel Fibres and Silica Fume

Mr. M. K. Maroliya

*Assistant professor, Applied Mechanics Dept, Faculty of Technology & Engineering,
M. S. University of Baroda, Vadodara,*

ABSTRACT: To focus on a structural performance in response to intense shear loading, a series of direct shear specimens random oriented fibres in shear failure plane were tested using optimized composition of RPC A simple shear strength test setup is proposed which is found to provide reliable and consistent results. The experimental results show that RPC exhibits ductile failure mode, higher ultimate strength and slip capacity in addition to much improved structural integrity. This enhancement of performance, however, reduces with decrease in size of Specimen.

Key Words: RPC, Strain hardening, Ductility, Slip capacity.

I. INTRODUCTION

Shear failure is generally brittle in concrete structures. Examples of concrete structural failure related to shear loading are bridge deck punching shear, corbel failure, anchor bolt pull out and segmental bridge shear key failure a motivation of the work presented here is to modify the brittle failure mode by taking advantage of unique material behavior of RPC. It is a fiber reinforced cement based material with designed microstructure such that composite undergoes pseudo strain hardening instead of softening when first crack strain is exceeded. As a result, RPC is characterized by ultimate tensile strain and fracture energies which can be higher than conventional cementitious materials. The present study is primarily concerned with the translation of pseudo strain hardening properties of RPC from material level to structural level.

A number of researchers have reported about the use of fibers as a replacement for shear reinforcement to enhance shear capacity of concrete beams in order to ensure a ductile final failure in flexure in RCC beams [3]. They emphasized testing of conventional reinforced beams to study the effect of shear span to effective depth ratio, fiber type, volume fraction and aspect ratio and the longitudinal steel content on shear capacity. But experience shows that two planes failing simultaneously in double shear for beam specimens never happens in reality and hence shear strength calculated in this manner could be erroneous. On the other hand, an attempt to get one failure plane in shear in beam and column specimen directly under compression testing machine invites undue eccentricity since one portion of the specimen needs to be held fixed with respect to the other part. Hence, there exists no standard, reliable and simple method to get direct shear strength. Sliding of one layer over the other layer with slip at common surface of contact gives realistic picture of direct shear. An attempt is made in the present paper to get direct shear strength of RPC material using a simplified and reliable test setup, both in conventional UTM under load control with large size specimen giving 2D effect of fiber orientation and MTS machine under displacement control with smaller size specimen giving 1D effect of randomly oriented fibres in shear failure plane.

Reactive Powder Concrete has higher durability, higher fatigue, and impact and abrasion resistances. With these unique properties, there is a potential for substantial reductions in cross-section dimensions reducing the weight of members which directly impact on substructure and erection cost. Further, for RPC beams without conventional shear reinforcement lower manpower and supervision cost are also a consideration when considering total long term costs. In term of structure performance, the inclusion of discrete fibers into the concrete matrix can arrest cracks and thereby control crack propagation.

II. MATERIAL COMPOSITION

TABLE: 1.0 OPTIMIZED COMPOSITION OF RPC

Mixture		RPC ₁	RPC ₂	RPCF ₃	RPCF ₄
Cement		1	1	1	1
Silica Fume		0.32	0.32	0.32	0.32
Quartz Sand		--	0.36	--	0.36
Sand	150 – 600µm	1.50	1.50	1.50	1.50
Super Plasticizer		0.035	0.032	0.032	0.032
Steel Fibers		--	--	0.20	0.20
Water/Cement		0.20	0.22	0.22	0.23
Compressive Strength (7-days) N/mm ²		96	118	103	138
Compressive Strength (28-days) N/mm ²		106	129	114	151
Flexural Strength (N/mm ²)		13.5	19.0	17.5	29

RPC₁ = RPC without Steel Fibers at normal Curing
 RPC₂ = RPC with Steel Fibers at hot water Curing
 RPCF₁ = RPC with Steel Fibers at normal Curing
 RPCF₂ = RPC with Steel Fibers at hot water Curing

III. LITERATURE REVIEW

Bairagi and Modhera [11, 12] checked the feasibility and reliability of the test method proposed by them with the test method suggested by JSCE method. Results obtained by their proposed method were 10% higher than that of JSCE method. They concluded that by comparing the procedure of test methods and fabrication of test assemblies as well as the test specimen, their proposed method is simpler to handle compared to JSCE method.

During the performance of the test using the arrangement as suggested by Bairagi and Modhera, it was seen that the roller penetrates down in the specimen, not giving the correct picture of the failure. Also the results reported by them were on higher side compared to that of JSCE method in case of fiber reinforced concrete samples.

IV. SPECIMEN AND TEST SET-UP

In the present study the influence of steel fiber variation on Shear strength of RPC have been studied. Different size of L-shape moulds are used for the preparation of samples. Three different percentage of steel fiber (i.e. 0.0, 1.0, 2.0 & 3.0%) added in the RPC mixture and compare the shear strength of RPC. The variations in the shear strength gained from the different steel fiber variation are tabulated in the table.

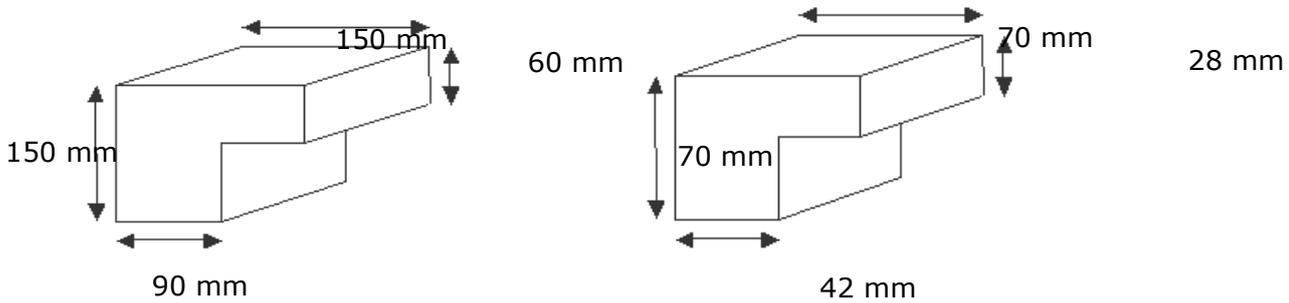


FIG.1.0: TWO TYPES OF L-SHAPED SPECIMENS



FIG.2.0: SHEAR TEST SET UP ON MTS

V. RESULTS AND DISCUSSION

Experimental results presented shows that the ultimate single shear strength for RPC with inclusion of fibers increases considerably with reference to plain RPC. With addition of fiber the specimens do not fail suddenly and the failure load is more than the first crack load. The quantity of fibers used in plain RPC do not significantly affect the first cracking load but a significantly influence on the rate of crack propagation and on the failure load. Typical shear strength variation with respect to fiber volume fraction and shear plane variations has been plotted in FIG 3.0 & FIG 4.0. It can be noted that the maximum increase in shear strength is found for 2% fiber volume fraction than 0%, 1%, 3% fiber volume.

For L shape shear specimen we had observed that shear strength increase for shear plane width 23mm than 28, 33, 37mm in 2% fiber volume in FIG 4.0 and shear plane 28, 33, 37mm of 2% fiber volume shear strength increase in compare shear plane 28, 33, 37mm of 0%, 1%, 3% fiber volume.

By referring fig.5 to fig.8 it is observed that the area below the curve is increases we increases fibre content as well as post peak performance is also enhanced to a large extant.

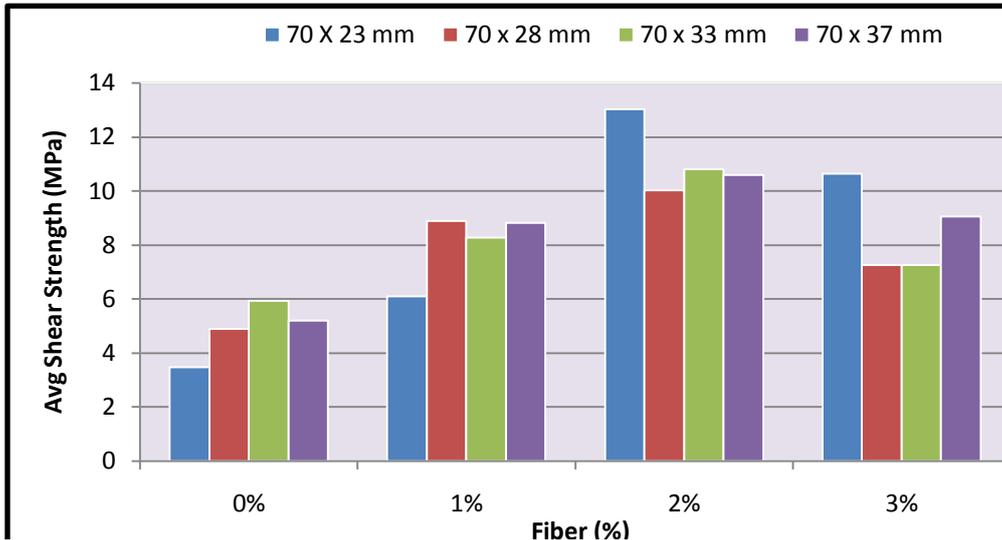


FIG.3.0: SHEAR STRENGTH OF RPC WITH STEEL FIBRE VARIATION

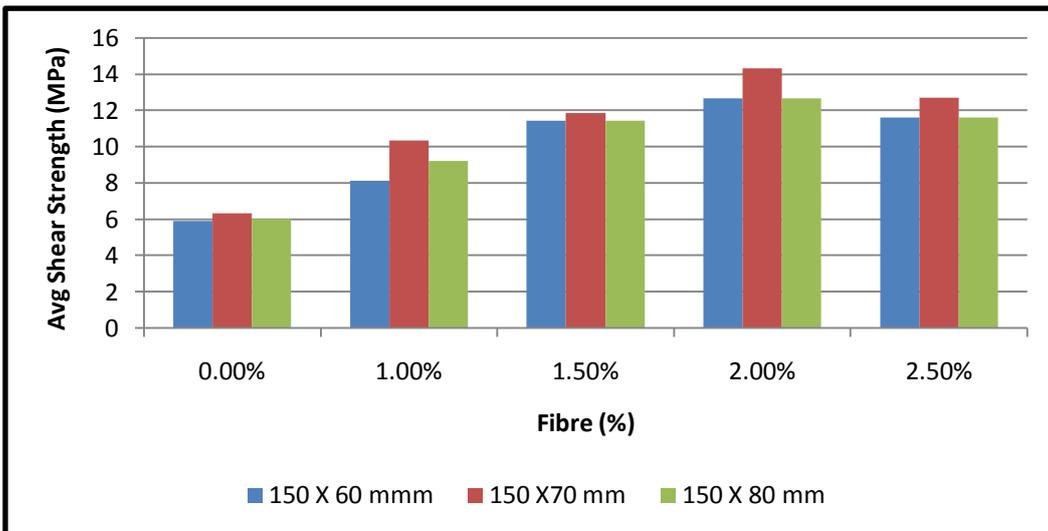


FIG: 4.0 SHEAR STRENGTH OF RPC WITH STEEL FIBRE VARIATIONS.

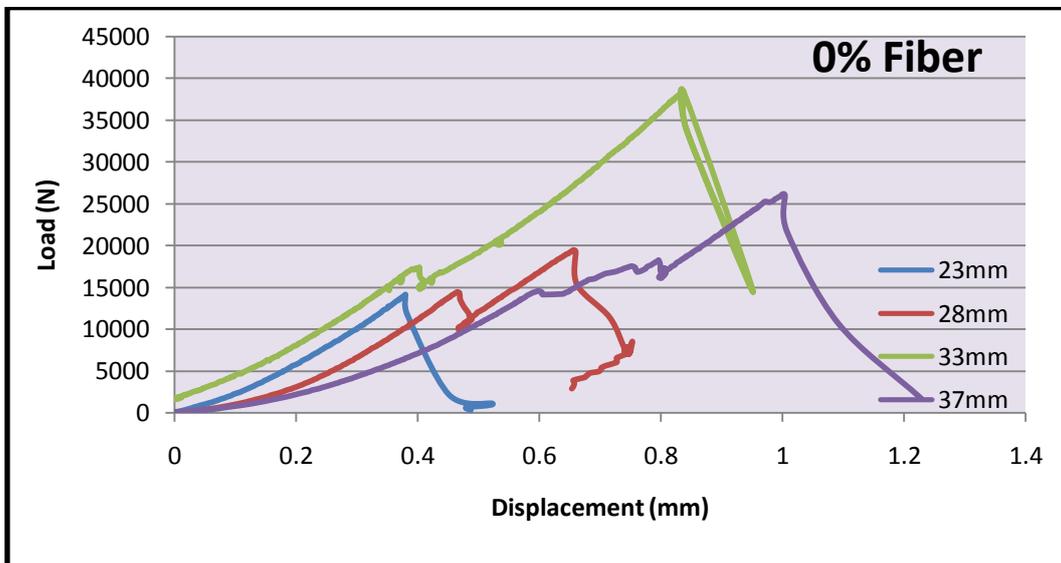


FIG.5.0: EFFECT OF 0.0% STEEL FIBER IN DIFFERENT SHEAR DEPTH

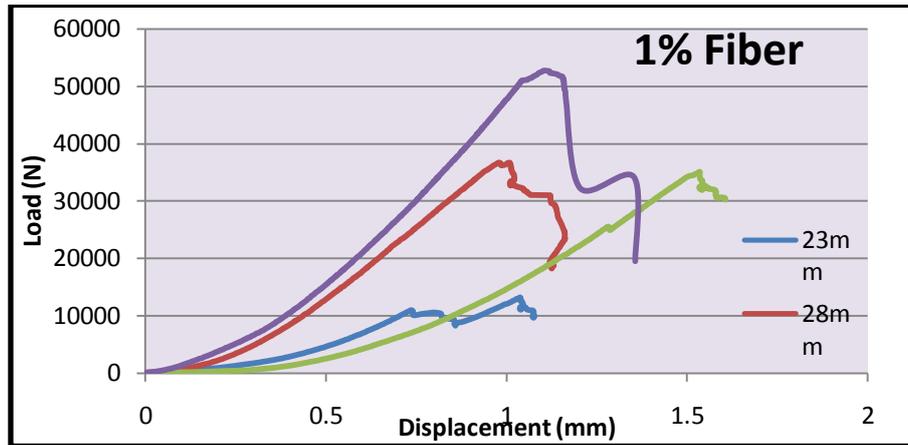


FIG.6.0: EFFECT OF 1.0% STEEL FIBER IN DIFFERENT SHEAR DEPTH

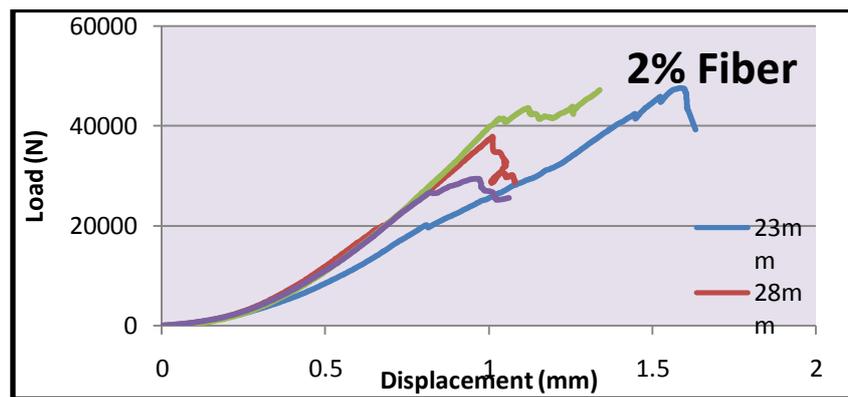


FIG.7.0: EFFECT OF 2.0% STEEL FIBER IN DIFFERENT SHEAR DEPTH

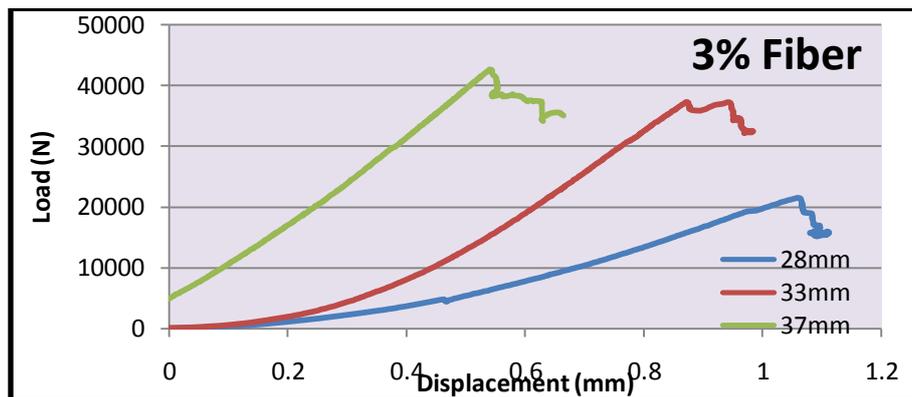


FIG.8.0: EFFECT OF 3.0% STEEL FIBER IN DIFFERENT SHEAR DEPTH

VI. CONCLUSIONS

- The proposed direct shear test is not only simple but also most suitable for the RPC composite. It indicates more realistic and efficient sliding mechanism in shear. Steel tends to enhance the shear strength much better than orientation effect could be closely seen as the shear strength of the smaller specimen and bigger specimens for all the matrices, deformation characteristics in shear could be evaluated better by testing the specimens in displacement control instead of load control.
- Maximum shear strength obtained in SF-RPC is quite high then ordinary concrete However, results of SF-RPC are much scattered as most of the specimens failed by fracture control after a formation of first crack, steel fibres being very strong and stiff in shear make it possible to resist large amount of load if fibres are available in the resisting plane.
- Due to presence of fibres ductile failure is observed .In fact crack get arrested by fibres does not get fractured but pulling out of fibres from matrix is observed.
- Experimental results shows that higher the thickness of shearing plane higher the value of shearing strength. However, optimum value results can be obtained at 60 to 75 mm depth.

- RPC specimens achieved much higher load carrying capacity and slip capacity in shear and were gradually damaged by ductile yielding. This phenomenon is due to ductile nature of RPC material and results of which confirms switching of failure mode from brittle concrete fracture to ductile yielding of RPC material with improved structural integrity. Such advantageous translation of structural shear response can effectively improve seismic resistance of building frames

REFERENCES

- 1) A. S. Dili and Manu Santhanam (2004) "*Investigation on Reactive Powder Concrete: A Developing Ultra High-Strength Technology*" Indian Concrete Journal.
- 2) Abdeldjelil and Thomas T.C.Hsu (1995) "*Constitutive Laws of Softened in Biaxial Tension-Compression*" ACI Structural Journal, Vol 92 No.[5]
- 3) Abouzar Sadrekarimi. (2004) "*Development Of A Light Weight Reactive Powder Concrete*". Journal Of Advanced Concrete Technology Vol. 2, no. 3, 409-417.
- 4) Aicha Kamen, Emmanuel Denarie And Eugen Bruhwiler. (2007) "*Thermal Effects On Physico-Mechanical Properties Of Ultra-High-Performance Fiber-Reinforced Concrete*", ACI Material Journal.
- 5) Ali R. Khaloo and Nakseok kim. (1997) "*Influence Of Concrete And Fiber Characteristics On Behaviour Of Steel Fiber Reinforced Concrete Under Direct Shear*" ACI Structural Journal, Vol. 94 No.[6].
- 6) Ammar Hassan And Makoto Kawakami (2005) "*Steel Free Composite Slabs Made Of Reactive Powder Materials And Fiber Reinforced Concrete*, ACI Structural Journal, Vol. 102 No. [5].
- 7) Arnaud Poitou, Francisco Chinesta, And Ge'ard Bernier. "*Orienting Fibers by Extrusion In Reinforced Reactive Powder Concrete*".
- 8) ASTM International Standard Test Method for *Compressive Strength of Hydraulic Cement Mortars*, (using [50-mm] cube specimens), C-109/c 109m-05, C-109-99.
- 9) ASTM International Standard Test Method for *Flow Of Hydraulic Cement Mortar*, C-1437-01,
- 10) B.I.G.BARR and K.L.W.LIU (1983) "*A compact test specimen*" Journal Of Materials Science Letters 2663-664.
- 11) Bairagi N. K. and Modhera C.D. (2001) "*Shear Strength of Fibre Reinforced Concrete*" ICI Journal Vol 1 No.[4].
- 12) Bairagi, n. K. & Modhera, C. D. (2004) "*An Experimental Study Of Shear Strength Test Method For SFRC*", Proceedings Of International Conference On Advances In Concrete And Construction, Vol. No. 1.

Cost Optimization Of Doubly Reinforced Rectangular Beam Section

S. A. Bhalchandra¹, P.K.Adsul²

^{1,2}(Applied Mechanics Department, Government College of Engineering Aurangabad, India)

ABSTRACT : In this paper, optimum design of simply supported doubly reinforced beams with uniformly distributed and concentrated load has been done by incorporating actual self weight of beam, parabolic stress block, moment-equilibrium and serviceability constraints besides other constraints. The optimization techniques in general enable designers to find the best design for the structure under consideration. In this particular case, the principal design objective is to minimize the total cost of a structure. The resulting structure, however, should not only be marked with a low price but also comply with all strength and serviceability requirements for a given level of the applied load. Total cost includes cost of concrete, cost of steel and cost of formwork are considered. A comparative study between the classical optimization techniques, namely the Generalized Reduced Gradient Method, Interior point algorithm optimization technique using MATLAB and one of the heuristic techniques, namely the Genetic Algorithm was carried out in this research. The initial solution for the optimization procedure has been obtained using limit state design as per IS: 456-2000.

Keywords: cost optimization, generalized reduced gradient method, genetic algorithm, interior point method, reinforced concrete.

I. Introduction

Optimum design of structures has been the topic of many studies in the field of structural design. A designer's goal is to develop an "optimal solution" for the structural design under consideration. An optimal solution normally implies the most economic structure without impairing the functional purposes the structure is supposed to serve.

Structural design requires judgment, intuition and experience, besides the ability to design structures to be safe, serviceable and economical. The design codes do not directly give a design satisfying all of the above conditions. Thus, a designer has to execute a number of design-analyze cycles before converging on the best solution. The intuitive design experience of an expert designer can give a good initial solution, which can reduce the number of design-analyze cycles. The optimization involves choosing of the design variables in such a way that the cost of the beam is the minimum, subject to the satisfaction of behavioral and geometrical constraints as per recommended method of design codes. Doubly reinforced beams (DRB) are required to be designed when the depth of the beam is restricted by architectural considerations and the beam has to take moment greater than limiting moment of resistance of the corresponding singly reinforced beam (SRB).

Some structure optimization work deals with minimization of cost of the structure using G.A. [1, 2, 3], some of the researchers have worked on cost optimization of the doubly reinforced concrete beam structure [4, 5, 6],

most of researchers have worked on cost optimization of reinforced concrete structure [7, 8], geometric programming model which gives the unique least-cost design of a beam, considering the cost of materials and shuttering and the structural requirements [9]. Whereas a natural velocity field method for shape optimization of reinforced concrete (RC) flexural members has been demonstrated [10]. The application of the Lagrangian Multiplier Method (LMM) to the minimum cost design of both singly and doubly reinforced concrete rectangular beams under limit state design conditions [11].

An initial solution for each case is obtained using the limit state method, by including self weight of the beam and considering parabolic stress block. The limit state design and the optimization is performed, subject to satisfaction of moment capacity, actual deflection and durability behavioral constraints, besides other geometrical constraints as recommended in IS: 456-2000 [12].

II. Optimization Technique

2.1 Classical Search and Optimization techniques

Traditional search and optimization methods can be classified into two distinct groups: Direct and gradient-based methods. In direct methods, only objective function and constraints are used to guide the search strategy, whereas gradient-based methods use the first and/or second-order derivatives of the objective function and/or constraints to guide the search process. Since derivative information is not used, the direct search methods are usually slow, requiring many function evaluations for convergence. For the same reason, they can be applied to many problems without a major change of the algorithm. On the other hand, gradient-based methods quickly converge to an optimal solution, but are not efficient in non differentiable or discontinuous problems.

2.1.1 The Generalized Reduced Gradient Method

The Generalized Reduced Gradient (GRG) Methods are algorithms for solving nonlinear programs of general structure.

GRG uses first partial derivatives of each function with respect to each variable. These are automatically computed by finite difference approximation (either forward or central differences). After an initial data entry segment, the program operates in two phases. If the initial values of the variables supplied by the user do not satisfy all the constraints, a Phase I optimization is started. The Phase I objective function is the sum of the constraint violations plus, optionally, a fraction of the true objective. This optimization terminates either with a message that the problem is infeasible or with a feasible solution. Beware if an infeasibility message is produced, because the program may have become stuck at a local minimum of the Phase I

objective (or too large a part of the true objective is incorporated), and the problem may actually have feasible solutions. The suggested remedy, in this case, is to choose different starting values for the variables (or reduce the proportion of the true objective) and try again.

Phase II begins with a feasible solution, either found by Phase I or with the user provided starting point if it is feasible, and attempts to optimize the objective function. At the conclusion of Phase II, a full optimization cycle has been completed and summary output is provided.

2.1.2 Interior Point Algorithm

Interior point methods (also referred to as barrier methods) are a certain class of algorithms to solve linear and nonlinear optimization problems. The interior point algorithm is used for general nonlinear optimization. It is especially useful for large-scale problems that have sparsity or structure, and tolerates user-defined objective and constraint function evaluation failures. It is based on a barrier function, and optionally keeps all iterates strictly feasible with respect to bounds during the optimization run.

In interior point method, the slack variables are introduced in to the simple non linear program, to make all inequality constraint in to non negativity, these non negativity constraints are replaced with logarithmic barrier terms in the objective. Incorporate the equality constraints into the objective using Lagrange multipliers. Newton's method is applied to compute search directions. Iterations are carried out and results are obtained using MATLAB.

2.2 Heuristic optimization techniques

In the last three decades, heuristic methods have been rapidly developed to solve optimization problems. These methods are principally intuitive and do not have theoretical support. Heuristic methods such as genetic algorithms (GAs), simulated annealing (SA) and tabu search (TS) provide general ways to search for a good but not necessarily the best solution.

2.2.1 Genetic Algorithm

Genetic algorithms (GA) are numerical optimization techniques inspired by the natural evolution laws. A GA starts searching design space with a population of designs, which are initially created over the design space at random. In the basic GA, every individual of population (design) is described by a binary string (encoded form). GA uses four main operators, namely, selection, creation of the mating pool, crossover and mutation to direct the population of designs towards the optimum design. In the selection process, some designs of a population are selected by randomized methods for GA operations, for example in creation of the mating pool, some good designs in the population is selected and copied to form a mating pool. The better (fitter) designs have a greater chance to be selected. Crossover allows the characteristics of the designs to be altered. In this process different digits of binary strings of each parent are transferred to their children (new designs produced by the crossover operation). Mutation is an occasional random change of the value of some randomly selected design variables. The mutation operation changes each bit of string from 0 to 1 or vice versa in a design's binary code depending on the mutation probability.

Mutation can be considered as a factor preventing from premature convergence.

III. Problem Formulation

The general form of an optimization problem is as follows

1. Given - Constant parameters
2. Find - Design variables
3. Minimize - Objective function
4. Satisfy - Design constraint

3.1 Constant Parameters

In this work, optimal design of doubly reinforced beam has been done for different material combinations of M20, M25 grades of concrete and Fe415, Fe500 grades of steel. The cost of materials for different grades and form work are given in Table 1.

Concrete Grade	C_c (Rs/m ³)	Steel Grade	C_s (Rs/kg)	C_f Rs/m ²
M20	4366	Fe415	58	320
M25	5610	Fe500	60	

3.2 Design variables

Width of beam = $b = x_1$

Tension reinforcement = $A_{st} = x_2$

Compression reinforcement = $A_{sc} = x_3$

Nominal cover = $d' = x_4$

3.3 Objective function

The objective function to be minimized:

$$F(x) = C_c [b (d + d') - (A_{st} + A_{sc})] + C_s [A_{st} + A_{sc}] + C_f [b + 2 (d + d')]$$

$$F(x) = C_c [X_1 (d + X_4) - (X_2 + X_3)] + C_s [X_2 + X_3] + C_f [X_1 + 2 (d + X_4)]$$

Where, C_c is cost of concrete, C_s is cost of steel and C_f is cost of formwork.

3.4 Design constraint

Geometrical constraints:

1. Ductility constraint: $x_u \leq x_a$

2. Constraint for minimum area of tension reinforcement:

$$A_{st} > \frac{0.85 b d}{f_y}$$

3. Constraint for maximum area of tension reinforcement:

$$A_{st} \leq 0.04 b D$$

4. Constraint for maximum area of compression reinforcement:

$$A_{sc} \leq 0.04 b D$$

5. Depth to width ratio constraint: $r = d/b$; $1.5 \leq r \leq 4$

Behavioral constraints:

1. Durability constraint: nominal cover ≥ 40 mm

2. Moment-equilibrium constraint: $M_u \leq M_c$

3. Deflection constraint (serviceability constraint): $\delta_{tot} \leq \delta_{all}$

Where, x_u is balance position of neutral axis in mm and x_a is actual position of neutral axis. M_u is bending moment

due to given loading and self weight in km.m, M_c is moment capacity of beam in kn.m. δ_{tot} and δ_{all} are sum of short term and long term deflection and allowable deflection in mm respectively.

IV. System Performance

In present study the attempt is made to optimize Doubly Reinforced Beam using Genetic Algorithm (MATLAB Toolbox) and performance analysis is performed of this element. Variation in parameters such as cost function, design variables are examined for several trial giving initial values obtained by Limit State Method. For justification of variations, Interior point Algorithm using Matlab and Generalized reduced gradient method (Microsoft Office Excel Solver Tool) are used.

The input of simply supported doubly reinforced beam (DRB) consists of 6 inputs, viz. load, span, d to b ratio, f_{ck} and f_y , and depth of beam. The output are cover to reinforcement (d') and optimum tensile steel reinforcement (pt). and optimum compression steel reinforcement (pc). In this paper five problems are solved by using three optimization techniques i.e. GRGM, IP and G.A., results are shown in Table 2.

4.1 Design example: S.S doubly reinforced beam with UDL

Design a simply supported doubly reinforced beam of span 9 m, depth 673.62 mm, subjected to following load and specification: Superimposed load = 10 kn/m, Live load = 20 kn/m, using M20 grade concrete and Fe415 grade steel.

4.1.1 Conventional limit state solution:

$M_u = 577.29$ kn/m, $P_t = 0.9789$, $P_c = 0.015$,

$M_c = 573.14$ kn/m, Cost = 2020.4 Rs/m.

4.1.2 Solution by proposed technique:

$M_u = 577.29$ kn/m, $P_t = 1.002$, $P_c = 0.04450$,

$M_c = 577.30$ kn/m, Cost = 1968.26 Rs/m.

Design given by proposed technique is safe while that given by conventional limit state method fails in moment capacity.

V. Conclusions

The main conclusions drawn from the current research are summarized as follows:

1. The results obtained from the Genetic Algorithm optimization technique showed a cost that is less than the cost obtained from the Generalized Reduced Gradient technique and Interior Point optimization technique. This comparison showed the superiority of the Genetic Algorithm technique over the classical Generalized Reduced Gradient technique and Interior Point optimization technique
2. It was shown that the Genetic Algorithm optimizer does a remarkable effort on minimizing the expensive material in the objective function of the numerical examples. This effort is devoted to the total cost. Therefore, one can conclude that the Genetic Algorithm search and optimization technique is powerful and intelligent
3. The performance of the Genetic Algorithm using different methods of crossover and selection can vary from one problem to another. Therefore, several values of each of the operators of the Genetic Algorithm should be examined in order to reach the best value for that operator

for the problem under consideration. This is called tuning of the Genetic Algorithm operators.

4. It can be said that researches carried out for finding optimum design of concrete structures are of great value to practicing engineers. The optimum solution satisfies the provisions of the code and minimizes the cost of the structure.

Additional research studies should be carried out on geometry and layout optimization of structural elements within the RC structure. Extra research should be carried out on other types of structures such as frames and trusses. The researcher recommends that there is a need for research on cost optimization of realistic RC three-dimensional large-scale structures.

REFERENCES

Journal Papers:

- [1] Saini B, Sehgal V.K. and Gambhir M.L., Genetically Optimized Artificial Neural Network Based Optimum Design Of Singly And Doubly Reinforced Concrete Beams, asian journal of civil engineering (building and housing), vol. 7, no. 6 (2006), pp 603-619
- [2] Coello C. and Farrea F.A., Use Of Genetic Algorithm For The Optimal Design Of Reinforced Concrete Beams.
- [3] Leps M. and Sejnoha M., New Approach To Optimization Of Reinforced Concrete Beams, Computers and Structures 81 (2003), pp 1957–1966, science direct.
- [4] Barros M.H.F.M., Martins R.A.F., Cost Optimization of singly and Doubly Reinforced Beams with EC2-2001, Struct. Multidisc. Optim., Springer-Verlag London limited, vol. 30,2005 , pp. 236-242.
- [5] Balaguru P., Cost Optimum Design Of Doubly Reinforced Concrete Beams, building and environment, vol 15, pp 219-222.
- [6] Prakash A., Agarwala S. K. and Singh K. K., Optimum Design Of Reinforced Concrete Sections, Computers and Structures Vol. 30. No. 4.
- [7] Kanagasundaram S. and Karihaloo B.L., Minimum cost design of reinforced concrete structures, structural optimization 2, pp.173-184.
- [8] Sarma K. C. and Adeli H, Cost Optimization Of Concrete Structures, Journal of Structural Engineering, Vol. 124, No.5, May, 1998.
- [9] Chakrabarty B. K., Models For Optimal Design Of Reinforced Concrete Beam, Computers & Structures Vol. 42, No. 3, pp. 447-451, 1992.
- [10] Rath D.P., Ahlawat A.S., and Ramaswamy A., Shape Optimization of RC Flexural Members, journal of structural Engineering ,ASCE, Vol. 125, No.2 , December 1999, pp. 1439-1446.
- [11] Ceranic B. and Fryer C., Sensitivity analysis and optimum design curves for the minimum cost design of singly and doubly reinforced concrete beams, Struct Multidisc Optim 20, pp 260–268.
- [12] IS 456-2000, Code of Practice for Plain and Reinforced Concrete, Bureau of Indian Standards, New Delhi.

- [13] Dr. Shah V. L. & Late Dr. Karve S. R., Limit State Theory and Design of Reinforced Concrete, Structures Publications, Year 2005, IV Edition, pp. 27-78,130-148.
- [14] Krisna Raju N., Design of Reinforced Concrete Structures, (IS 456-2000) III Edition CBS Publishers.
- [15] Dr. Syal IC & Dr. Goel A.K., Reinforced Concrete Structures, S. Chand & Company.
- [16] Rao S.S., Engineering Optimization Theory and Practice, New Age International Publisher, 2006, III Edition, pp. 29-33.

Table 2: Results for optimal design for SS - Doubly Reinforced Beam

Sr	Input	Method	Pt	Pc	Mu	Mc	F(x)
1	W=350 kN L=4 m, r=1.5 M20, Fe415 d=660.07mm	GRGM	0.98655	0.02967	548	548.09	1950.93
		I.P	0.9904	0.03628	548	548.176	1942.8
		G.A.	1.03281	0.07377	548	555.23	1888.475
2	W=40 Kn/m L=10m, r=1.5 M20, Fe500 d=781.38mm	GRGM	0.8367	0.08094	950.0	950.49	2588.98
		I.P	0.8404	0.08385	950.0	951.44	2577.35
		G.A	0.8484	0.08741	950.0	950.51	2549.12
3	W=30 Kn/m L = 9 m, r=1.5 M20, Fe415 d=673.62mm	GRGM	0.9789	0.0218	577.29	577.28	2017.11
		I.P	0.9967	0.0367	577.29	577.90	1989.522
		G.A	1.002	0.04450	577.29	577.30	1968.26

Resource and Quality Aware Location Anonymization Mechanism For Wireless Sensor Networks

Ch. Swarna Latha¹, D.T.V. Dharmajee Rao²

(M.Tech, Department of CSE, Aditya Institute of Technology and Management, Andhra Pradesh, India)

(Professor & HOD, Department of CSE, Aditya Institute of Technology and Management, Andhra Pradesh, India)

ABSTRACT: A wireless sensor network is a heterogeneous network consisting of a large number of tiny low-cost nodes and one or more base stations. These networks can use in various applications like military, health and commercial. However, the privacy preservation problem has drawn huge attention in the research community. This problem is exacerbated in the domain of WSNs due to the extreme resource limitation of sensor nodes. In this paper, we proposed a model for privacy preservation for mobile users by using anonymization and aggregate location monitoring in a wireless sensor network. Resource-aware and quality-aware anonymization algorithms are designed to preserve personal location and provide location monitoring services. Sensor nodes execute location anonymization algorithms to provide k -anonymous aggregate locations. To evaluate the performance of the proposed algorithm we simulate it in the NS2 simulator. Our experimental results show that proposed solution provides high quality location monitoring services for end users and guarantees the location privacy of the monitored persons.

Keywords: Aggregate location, Anonymization, Cloaked area, Sensor node, WSN, NS2.

I. Introduction

A wireless sensor network (WSN) is a heterogeneous network consisting of a large number of tiny low-cost nodes (devices) and one or more base stations (sinks) [1]. Main purpose of the WSN is to monitor some physical phenomena (e.g., temperature, barometric pressure, light) inside an area of deployment. Nodes are equipped with radio transceiver, processing unit, battery and sensor(s). Nodes are constrained in processing power and energy, whereas the base stations are not severely energy resources. The base station act as gateways between the WSN and other networks such as Internet etc.. The WSN is used in various applications like military, health and commercial. WSNs are becoming one of the building blocks of pervasive computing. They provide simple and cheap mechanism for monitoring in the specified area. But WSN technology is an inappropriate use can significantly violate privacy of humans. WSNs are frequently deployed to collect sensitive information. WSN can be used to monitor the movements of traffic in a city. Such a network can be used to determine location of people or vehicles [2].

If this information is available on a wide basis it can easily lead to blackmailing or stalking. It can be also exploited by terrorists as a targeting tool to impact specific people or buildings. Another example of a WSN application, in which privacy is heavily exposed, is health monitoring. Here, the medical measurements should be available only to the attending physician [4,5]. Wrong usage

of simple commercial WSNs can easily result into serious privacy violations as well. Suppose that the WSN monitors people movements at a supermarket to improve the placement of products within the shelves. If someone is able to find out detailed information related to a particular person, then a seemingly innocent application turns into a privacy violating tracking device. This example demonstrates that in most cases collected data they do not pose a privacy threat. The problem arises when the data can be linked to a specific person. This is why anonymity and proper identity management of the nodes, or their carriers, or the subjects that these nodes monitor, are needed. If an attacker is not able to link measured data with the measuring device or location then this data is of a little value for privacy attacks [6].

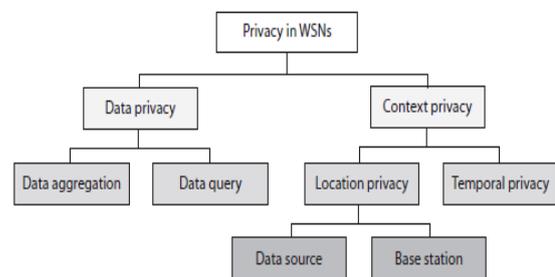


Fig 1: privacy preserving protections in WSNs

Privacy in the WSN is classified into Data-oriented and Context-oriented [3]. Data-oriented protections are then categorized into data aggregation and data query techniques. Context-oriented privacy protections can be split into location privacy and temporal privacy techniques, the location privacy is split into data source and base station techniques which is described in figure 1.

In this paper, we propose a privacy preservation of such mobile users with the help of anonymization and by reporting aggregate location. An anonymization means a person is indistinguishable amongst k persons in a network. The most effective way to compromise location privacy used by adversary is packet-tracing. In such an attack, an adversary can locate the immediate nodes by eavesdropping the transmitted packet, and further reduce the flow direction of packets. Even worse, the attacker can trace hop-by-hop towards the sink or source nodes. To defend against packet-tracing attack, many approaches are proposed. One of the approaches is providing aggregate location of a user. Along with privacy preservation of mobile users we are monitoring location of any mobile user through our system. Location monitoring is defined as monitoring every action, movement of any mobile user without disturbing its privacy.

The rest of the paper is organized as section 2: discuss about the related work, section 3: presents the Proposed Model, section 4: discuss about Proposed

Solution, section 5: discuss about Experimental setup, section 6: concludes the paper.

II. Related Work

Chaum [11] has started developing solutions for anonymous communications to provide privacy in WSNs. It is used to provide users with an anonymous e-mail system based on a special type of device called the mix. The main functionality of a mix is to receive a cryptographically encrypted message and transform it into a new message indistinguishable from the originally input one. In order to send a message, the source creates several layers of encryption over the message using the public keys of the different mixes that the message will traverse. Onion routing [9] and Tor [7] provide application independent anonymous connections in near real time by creating connections through a set of machines called the onion routers. Whenever an application establishes a connection, it first connects to an onion proxy, which is the entrance point to the anonymous network. The onion proxy is in charge of determining a series of onion routers that will define the bidirectional path that the packets of that specific connection will traverse. The path is constructed by using the cryptographic material of each of the onion routers, which is included in a data structure called the onion. Once the path has been established, the application data is sent through the onion network by adding a layer of encryption for each of the hops in the anonymous path. Each of the onion routers peels of its corresponding layer, changing the appearance of the data, and forwards it to the next onion router. The main drawback of this technique is based on a network core which the users must fully trust. Later Crowds [10] and Hordes [8] were proposed decentralized approaches. Both approaches are based on the idea of making individuals disappear into a group of peers. Upon receiving a message from a peer, the recipient will randomly choose whether to forward it to another peer or to finally submit it to the real destination. Each member of the path must remember its predecessor and successor so that subsequent messages coming from the same source follow the same path through the anonymous network. Note that any member of the path has only a local view of the route that a message traverses so that no peer can determine who the actual origin of a message is. Furthermore, since all communications are re-encrypted at every hop, a local eavesdropper cannot easily determine the destination of a message unless the originator decides to send the message directly to the destination. The main difference between Crowds and Hordes is in the way responses are sent back to the origin. In Hordes it is done by multicasting messages, which provides a better performance.

III. Proposed Model

The proposed Architecture consists of user, server and trusted zone and Sensor node, mobile users in a trusted zone. Anonymity level is set by administrator of a system to provide security for mobile users in a trusted zone. The mobility objects are shown in figure 2 by green color. If a user asks query regarding any user in a zone to a server then server passes this query to a sensor nodes present in trusted zone. Then sensor node from one area will exchange message with the other and report an aggregate location to the server and then server will send the answer to the user.

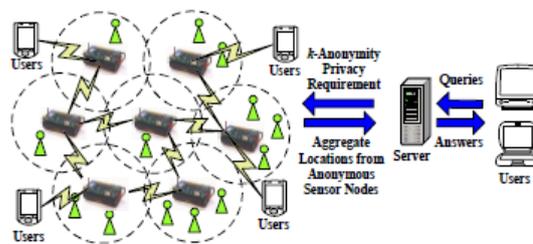


Fig 2: Proposed Architecture

IV. Proposed Solution

In our proposed solution we propose two algorithms

4.1 Resource aware algorithm

The main idea of the Resource aware algorithm is to find adequate number of persons in that network and accordingly finding a cloaked area as MBR (minimum bounded area).

Broadcast step

In this step, every sensor node in a network broadcasts a message which contains id, area and number of nodes to its nearest neighbor. In this way every sensor node forms its own table and also checks for adequate number of objects in its sensing area and accordingly it sends notification message to the nearer sensor nodes and follows the next step.

Cloaked area step

The basic idea of this step is that each sensor node blurs its sensing area into a cloaked area that includes at least k objects, in order to satisfy the k -anonymity privacy requirement. To minimize computational cost, it uses a greedy approach to find a cloaked area based on the information stored in table. Each sensor node initializes a set S and then determines a score for each peer in its table. The score is defined as a ratio of the object count of the peer to the distance between the peer and node. The score is calculated to select a set of peers from table to S to form a cloaked area that includes at least k objects and has an area as small as possible. Then we repeatedly select the peer with the highest score from the table to S until S contains at least k objects. Finally, node determines the cloaked area that is a minimum bounding rectangle that covers the sensing area of the sensor nodes in S , and the total number of objects in S .

Validation step:

Validation step is used to avoid reporting aggregate locations with a containment relationship to the server. We do not allow the sensor nodes to report their aggregate locations with the containment relationship to the server, because combining these aggregate locations may pose privacy leakage.

4.2 Quality aware algorithm

The quality-aware algorithm starts from a cloaked area A , which is computed by resource aware algorithm. Then A will be iteratively updated based on extra communication among the sensor nodes until its area reaches the minimal possible size. For both algorithms, the sensor node reports its cloaked area with the number of monitored persons in the area as an aggregate location to the server.

Search space step

Sensor network has a large number of sensor nodes hence it is very costly for a sensor node to gather the information of all the sensor nodes to compute its minimal cloaked area. To reduce the cost, node determines a search space based on the input cloaked area computed by the resource-aware algorithm.

The Minimal Cloaked Area step

This step takes a set of peers residing in the search space, S , as an input and computes the minimal cloaked area for the sensor node m . The basic idea of the first optimization technique is that we do not need to examine all the combinations of the peers in S , instead we only need to consider the combinations of at most four peers. Because at most two sensor nodes defines width of MBR and at most two sensor nodes defines height of MBR. It reduces cost by reducing the number of MBR computations among the peers in S . The second optimization technique has two properties, lattice structure and monotonicity property. In a lattice structure, a data set that contains n items can generate 2^{n-1} item sets excluding a null set. We generate the lattice structure from the lowest level based on a simple generation rule. The monotonicity property of a function f indicates that if X is a subset of Y , then $f(X)$ must not exceed $f(Y)$.

For our problem, the MBR of a set of sensor nodes S has the monotonicity property, because adding sensor nodes to S must not decrease the area of the MBR of S or the number of objects within the MBR of S .

The validation step

This step is to avoid reporting aggregate locations with a containment relationship to the server. We do not allow the sensor nodes to report their aggregate locations with the containment relationship to the server, because combining these aggregate locations may pose privacy leakage.

V. Experimental Setup

We have implemented our proposed algorithm in NS2, which has been highly validated by the networking research community. The simulation parameters where listed in table 1.

Attack model error: This metric measures the resilience of our system to the attacker model by the relative error between the estimated number of objects $b N^{\wedge}$ in a sensor node's sensing area and the actual one N .

Table 1: NS2 parameters

Parameters	Value
MAC Layer	IEEE 802.11
Number of nodes	20
Data rate	11Mbps
Packet Size	512 B
Simulation Duration	200 sec
Traffic Flow	TCP

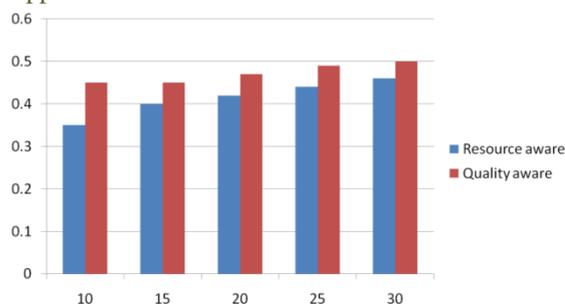


Fig 3: Attack model error vs anonymity levels

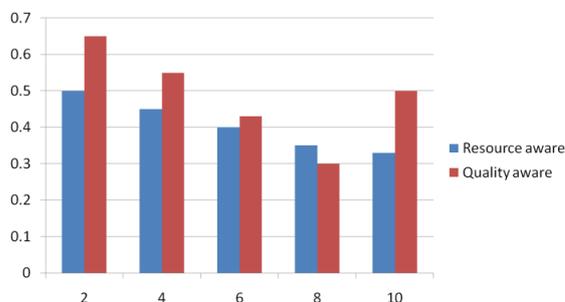


Fig 4: Attack model error vs number of objects in (thousands)

Figure 3 depicts that the stricter the anonymity level, the larger the attacker model error will be encountered by an adversary. When the anonymity level gets stricter, our algorithms generate larger cloaked areas, which reduce the accuracy of the aggregate locations reported to the server. Figure 4 shows that the attacker model error reduces, as the number of objects gets larger. This is because when there are more objects, our algorithms generate smaller cloaked areas, which increase the accuracy of the aggregate locations reported to the server. It is difficult to set a hard quantitative threshold for the attacker model error.

VI. Conclusion

A wireless sensor network is a heterogeneous network consisting of a large number of tiny low-cost nodes and one or more base stations. The privacy preservation problem has drawn huge attention in the research community. This problem is exacerbated in the domain of WSNs due to the extreme resource limitation of sensor nodes. To overcome privacy problem in WSN, we proposed a model for privacy preservation for mobile users by using anonymization and aggregate location monitoring in a wireless sensor network. Resource-aware and quality-aware anonymization algorithms are designed to preserve personal location and provide location monitoring services. Sensor nodes execute location anonymization algorithms to provide k -anonymous aggregate locations. To evaluate the performance of the proposed algorithm we simulate it in the NS2 simulator. Our experimental results show that proposed solution provides high quality location monitoring services for end users and guarantees the location privacy of the monitored persons.

References

- [1] D. Culler and M. S. Deborah Estrin, .Overview of sensor networks, IEEE Computer, 2004.

- [2] M. Gruteser, G. Schelle, A. Jain, R. Han, and Grunwald, Privacy-aware location sensor networks, 2003.
- [3] W. He, X.Liu, H. Nguyen, K. Nahrstedt and Abdelzaher, PDA: Privacy-preserving data aggregation in wireless sensor networks, IEEE, 2007.
- [4] C.-Y. Chow, M. F. Mokbel, and X. Liu, A peer-to-peer spatial cloaking algorithm for anonymous location-based services, In Proc. of ACM GIS, 2006.
- [5] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias "Preventing location-based identity inference in anonymous spatial queries" IEEE, 2007.
- [6] B.Son, S. Shin, J.Kim and Y.Her "Implementation of the Real-Time People Counting System using Wireless Sensor Networks," IJMUE, 2007.
- [7] R. Dingedine, N. Mathewson, and P. Syverson. Tor: The Second-Generation Onion Router. In SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium, pages 21{21, Berkeley, CA, USA, 2004.
- [8] B. N. Levine and C. Shields. Hordes: "A Multicast Based Protocol for Anonymity". IEEE International Conference, 2002.
- [9] M. Reed, P. Syverson, and D. Goldschlag. Anonymous Connections and Onion Routing. Selected Areas in Communications, IEEE Journal on, May 1998.
- [10] M. Reiter and A. Rubin. Crowds: Anonymity for Web Transactions. ACM transactions on information and system security, 1998.
- [11] Y. Ouyang, Z. Le, Y. Xu, N. Triandopoulos, S. Zhang, J. Ford, and F. Makedon. "Providing Anonymity in Wireless Sensor Networks", IEEE International Conference July 2007.

Fault Allocation in Multilevel Converter Statcoms

B.Bansiraj¹, M. Lokya²

¹Final M.Tech student, Dept of EEE, MIST, Sathupally, India,

²Asst. Professor, Dept of EEE, MIST, Sathupally, India,

Abstract: This paper introduces an approach to detect the existence of the faulted switch, identify which switch is faulty, and reconfigure the STATCOM. This approach is illustrated on an eleven-level STATCOM and the effect on the dynamic performance and the total harmonic distortion (THD) is analyzed. Many static synchronous compensators (STATCOMs) utilize multilevel converters due to the following: 1) lower harmonic injection into the power system; 2) decreased stress on the electronic components due to decreased voltages; and 3) lower switching losses. One disadvantage, however, is the increased likelihood of a switch failure due to the increased number of switches in a multilevel converter. Due to its unique topology, the Modular Multilevel STATCOM has many advantages but requires a sophisticated controller and puts higher requirements on simulation tools. To simulate the STATCOM in real-time is preferable because it enables hardware-in-the-loop test of the system in various scenarios including extreme fault conditions, which cannot be tested on a real STATCOM. A single switch failure, however, does not necessarily force an $(2n + 1)$ -level STATCOM offline. Even with a reduced number of switches, a STATCOM can still provide a significant range of control by removing the module of the faulted switch and continuing with $(2n - 1)$ levels.

Index Terms: Fault detection, multilevel converter, static synchronous compensator (STATCOM)

I. INTRODUCTION

THE static synchronous compensator (STATCOM) has been well accepted as a power system controller for improving voltage regulation and reactive compensation [1]–[5]. There are several compelling reasons to consider a multilevel converter topology for the STATCOM [6]–[8]. These well-known reasons include the following: 1) lower harmonic injection into the power system; 2) decreased stress on the electronic components due to decreased voltages; and 3) lower switching losses [9]. Various multilevel converters also readily lend themselves to a variety of PWM strategies to improve efficiency and control. An eleven-level cascaded multilevel STATCOM is shown in Fig. 1. This converter uses several full bridges in series to synthesize staircase waveforms. Because every full bridge can have three output voltages with different switching combinations, the number of output voltage levels is $2n + 1$ where n is the number of full bridges in every phase. The converter cells are identical and therefore modular.

As higher level converters are used for high output rating power applications, a large number of power switching devices will be used. Each of these devices is a potential failure point. Therefore, it is important to design a sophisticated control to produce a fault-tolerant STATCOM. A faulty power cell in a cascaded H-Bridge STATCOM can potentially cause switch modules to explode [10] leading to the fault conditions such as a short circuit or an overvoltage on the power system resulting in an expensive down time [11]. Subsequently, it is crucial to identify the existence and location of the fault for it to be removed. Several fault detection methods have been proposed over the last few years [10]–[18]. Resistor sensing, current transformation, and V_{CE} sensing are some of the more common approaches. For example, a method based on the output current behavior is used to identify IGBT short circuits [12]. The primary drawback with the proposed approach is that the fault detection time depends on the time constant of the load. Therefore, for loads with a large RL time constant, the faulty power cell can go undetected for numerous cycles, potentially leading to circuit damage. Another fault detection approach proposed in [13] is based on a switching frequency analysis of the output phase voltage. This method was applied to flying capacitor converters and has not been extended to cascaded converters. AI-based methods were proposed to extract pertinent signal features to detect faults in [14]. In [15], sensors are used to measure each IGBT current and to initiate switching if a fault is detected.

A fault-tolerant neutral point-clamped converter was proposed in [16]. In [17], a reconfiguration system based on bidirectional switches has been designed for three-phase asymmetric cascaded H-bridge inverters. The fundamental output voltage phase shifts are used to rebalance a faulted multilevel cascaded converter in [18].

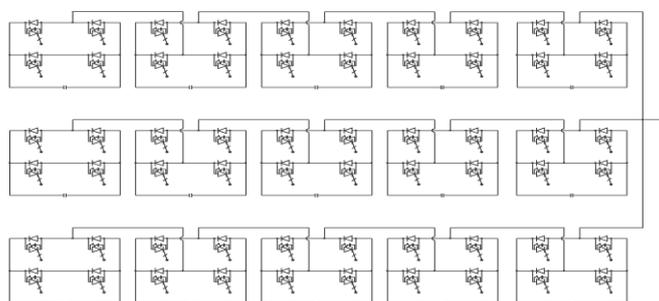


Fig. 1. Eleven-level cascaded multilevel STATCOM.

In this paper, the method we propose requires only that the output dc link voltage of each phase be measured. This measurement is typically accomplished anyway for control purposes. If a fault is detected, the module in which the fault occurred is then isolated and removed from service. This approach is consistent with the modular design of cascaded converters in which the cells are designed to be interchangeable and rapidly removed and replaced. Until the module is replaced, the multilevel STATCOM continues to operate with slightly decreased, but still acceptable, performance. In summary, this approach offers the following advantages:

- No additional sensing requirements;
- Additional hardware is limited to two by-pass switches per module;
- The dynamic performance and THD of the STATCOM is not significantly impacted.

II. MULTILEVEL STATCOM

A cascaded multilevel STATCOM contains several H-bridges in series to synthesize a staircase waveform. The inverter legs are identical and are therefore modular. In the eleven-level STATCOM, each leg has five H-bridges. Since each full bridge generates three different level voltages (V , 0 , $-V$) under different switching states, the number of output voltage levels will be eleven. A multilevel configuration offers several advantages over other converter types [19].

- 1) It is better suited for high-voltage, high-power applications than the conventional converters since the currents and voltages across the individual switching devices are smaller.
- 2) It generates a multistep staircase voltage waveform approaching a more sinusoidal output voltage by increasing the number of levels.
- 3) It has better dc voltage balancing, since each bridge has its own dc source.

To achieve a high-quality output voltage waveform, the voltages across all of the dc capacitors should maintain a constant value. Variations in load cause the dc capacitors to charge and discharge unevenly leading to different voltages in each leg of each phase. However, because of the redundancy in switching states, there is frequently more than one state that can synthesize any given voltage level. Therefore, there exists a “best” state among all the possible states that produces the most balanced voltages [20]. Since there are multiple possible switching states that can be used to synthesize a given voltage level, the particular switching topology is chosen such that the capacitors with the lowest voltages are charged or conversely, the capacitors with the highest voltages are discharged. This redundant state selection approach is used to maintain the total dc link voltage to a near constant value and each individual cell capacitor within a tight bound.

Different pulse width modulation (PWM) techniques have been used to obtain the multilevel converter output voltage. One common PWM approach is the phase shift PWM (PSPWM) switching concept [21]. The PSPWM strategy causes cancellation of all carrier and associated sideband harmonics up to the $(N - 1)$ th carrier group for an N -level converter. Each carrier signal is phase shifted by

$$\Delta\phi = \frac{2\pi}{n}$$

Where n is the number of cells in each phase. Fig. 2 illustrates the carrier and reference waveforms for a phase leg of the eleven-level STATCOM. In this figure, the carrier frequency has been decreased for better clarity. Normally, the carrier frequency for PWM is in the range of 1–10 kHz.

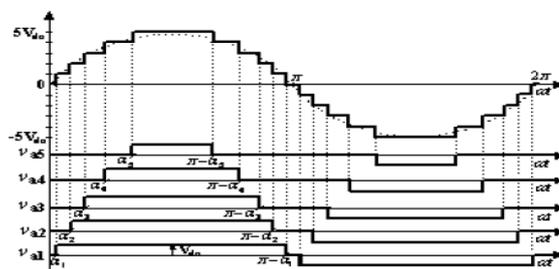


Fig. 2. (a) Carrier and reference waveform for PSPWM. (b) Output waveform.

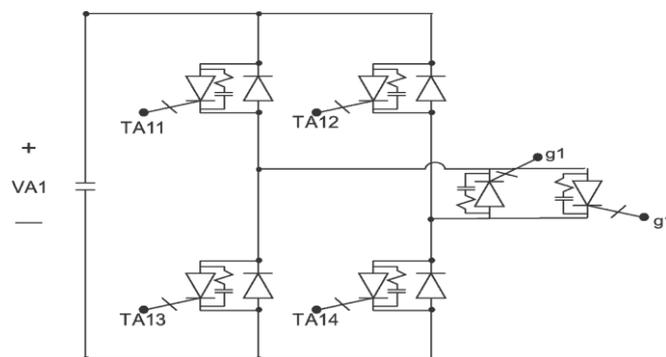


Fig. 3. Cell with fault switch.

III. ANALYSIS OF FAULT FOR THE MULTILEVEL STATCOM

A converter cell block, as shown in Fig. 3, can experience several types of faults. Each switch in the cell can fail in an open or closed state. The closed state is the most severe failure since it may lead to shoot through and short circuit the entire cell. An open circuit can be avoided by using a proper gate circuit to control the gate current of the switch during the failure [23]. If a short circuit failure occurs, the capacitors will rapidly discharge through the conducting switch pair if no protective action is taken. Hence, the counterpart switch to the failed switch must be quickly turned off to avoid system collapse due to a sharp current surge. Nomenclature for the proposed method is given in Table I.

The staircase voltage waveform shown in Fig. 2 is synthesized by combining the voltages of the various cells into the desired level of output voltage. At the middle levels of the voltage waveform, due to the switching state redundancy, there are more than one set of switching combinations that may be used to construct the desired voltage level. Therefore, by varying the switching patterns, the loss of any individual cell

TABLE I
 NOMENCLATURE

E_{out}	STATCOM output voltage (V)
\hat{E}_{out}	Filtered STATCOM output voltage (RMS) (V)
E'	STATCOM threshold voltage (constant) (V)
S_{j1}, S_{j2}	Switching signal of the j -th cell (0, 1)
J_i	possible STATCOM output voltage (V)
x_i	difference between possible and actual STATCOM output (V)
g_j	bypass signal for j -th cell (0, 1)

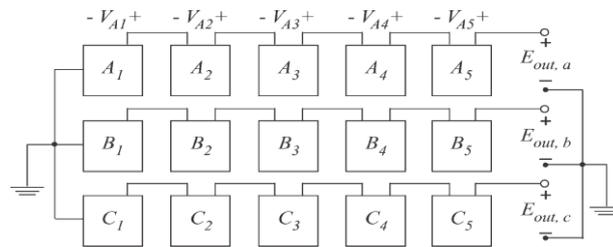


Fig. 4. Simplified eleven-level cascaded multilevel STATCOM.

will not significantly impact the middle voltages of the output voltage. However, the peak voltages require that all cells contribute to the voltage; therefore, the short circuit failure of any one cell will lead to the loss of the first and $(2n + 1)$ output levels and cause degradation in the ability of the STATCOM to produce the full output voltage level.

Consider the simplified eleven-level converter shown in Fig. 4. The process for identifying and removing the faulty cell block is summarized in Fig. 5. The input to the detection algorithm is \hat{E}_{out} for each phase, where \hat{E}_{out} is the STATCOM-filtered RMS output voltage. If the STATCOM RMS output voltage drops below a preset threshold value (E), then, a fault is known to have occurred (see Fig. 6).

Once a fault has been detected to have occurred, then, the next step is to identify the faulty cell. By utilizing the switching signals in each converter cell, (i.e., S_1 and S_2), it is possible to calculate all of the possible voltages that can be produced at any given instant as illustrated in Table II (terminology adopted from [23]): Thus, the output voltage of a cell is

$$v_{ax} = v_{ax} + -v_{ax} \quad (1)$$

and since the cells of the STATCOM are serially connected, the total output voltage per phase is

$$v_{y0} = \sum_{x=1}^n v_{yx}, \quad y \in [a, b, c]$$

where n is the number of blocks. By utilizing the switching signals in each converter cell, (i.e., S_{j1} and S_{j2} , j is the cell number), it is possible to calculate all of the possible voltages that can be produced at any given instant. When there is a fault in the multilevel converter, the capacitor at the faulty block will rapidly discharge. This discharge results in a phase shift in the output ac voltage as well as a change in amplitude of voltage.

TABLE II
 SWITCHING STATE AND OUTPUT VOLTAGE OF AN H-BRIDGE

S_1	S_2	v_{ax}^+	v_{ax}^-	v_{ax}
0	0	0	0	0
0	1	0	v_{dc}	$-v_{dc}$
1	0	v_{dc}	0	v_{dc}
1	1	v_{dc}	v_{dc}	0

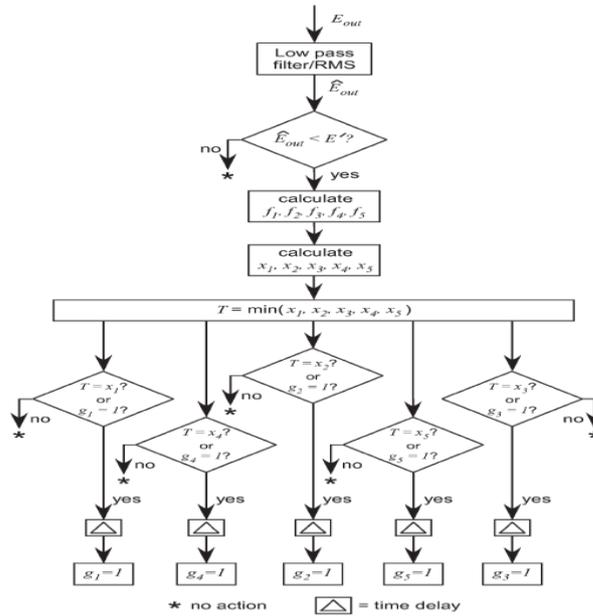


Fig. 5. Flowchart for eleven-level converter.

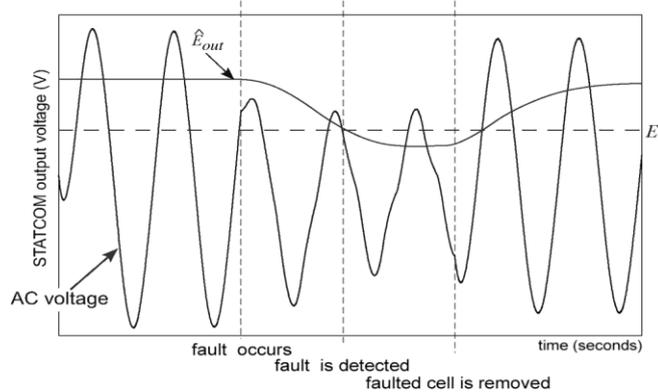


Fig. 6. STATCOM-filtered output voltage and threshold value.

The set of all possible phase fault voltages for an eleven-level converter is given by

$$f_1 = V_{dc0} (S_{21} - S_{22} + S_{31} - S_{32} + S_{41} - S_{42} + S_{51} - S_{52}) \text{ (cell 1 faulted)}$$

$$f_2 = V_{dc0} (S_{11} - S_{12} + S_{31} - S_{32} + S_{41} - S_{42} + S_{51} - S_{52}) \text{ (cell 2 faulted)}$$

$$\dots$$

$$f_5 = V_{dc0} (S_{11} - S_{12} + S_{21} - S_{22} + S_{31} - S_{32} + S_{41} - S_{42}) \text{ (cell 5 faulted)}$$

or more succinctly as

$$f_i = V_{dc0} \sum_{\substack{j=1 \\ j \neq i}}^n (S_{j1} - S_{j2}), \quad i = 1, \dots, n \quad (3)$$

where V_{dc0} is the ideal voltage across a single cell block. If there is a faulted cell, only one f_i will be near the actual STATCOM output phase voltage E_{out} ; all of the others will be too high. Therefore, to determine the location of the fault cell, each f_i is compared against E_{out} to yield

$$x_i = |E_{out} - f_i|, \quad i = 1, \dots, n. \quad (4)$$

The smallest x_i indicates the location of the faulted block because this indicates the f_i which most closely predicts the actual E_{out} . The choice of threshold voltage E depends on the number of cells in the converter. The ideal output voltage is

$$\hat{E}_{out,0} = \frac{nV_{dc0}}{\sqrt{2}}.$$

During a fault, E_{out} will decrease by V_{dc0} yielding

$$\hat{E}_{out,fault} = \frac{(n-1)V_{dc0}}{\sqrt{2}} = \frac{n-1}{n} \hat{E}_{out,0}.$$

Therefore, the threshold voltage E should be chosen such that $(n-1/n)E_{out,0} \leq E \leq E_{out,0}$. In an eleven-level converter, $n = 5$ and the faulted RMS voltage will decrease by roughly 20%. Therefore, a good choice for E is 85%

of the rated output STATCOM voltage.

The last step is to actuate the module bypass switch g_i shown in Fig. 3. A slight time delay is added to the logic to neglect for momentary spikes that may occur. It is desirable to neglect momentary sags in the dc link voltage, but respond to sags of increased duration that indicate a faulted module. Fig. 7 shows the realization logic for the proposed fault detection and module removal method.

The use of a fault handling switch in multilevel converters is not uncommon. In a fault handling switch is used in a fly- ing capacitor multilevel inverter. While the additional circuitry does increase the cost of the circuit, it also increases the relia- bility be enabling the circuit to keep working (albeit at a slightly reduced operating range) until the module can be replaced.

IV. RESULTS

The single line diagram of the electrical distribution system feeding an arc furnace is shown in Fig. 8. The STATCOM has been shown to be an efficient controller to mitigate arc furnace flicker [27]. The electrical network consists of a 115-kV generator and an impedance that is equivalent to that of a large network at the point of common coupling (PCC). The STATCOM is connected to the system through a Y-Delta transformer. The system was simulated using PSCAD/EMDTC. The electrical arc furnace load is non sinusoidal, unbalanced, and randomly fluctuating. Electric arc furnaces are typically used to melt steel and will produce current harmonics that are random. In addition to the integer harmonics, arc furnace currents are rich in inter harmonics [24]. The flicker waveform has sub synchronous variations in the 5–35-Hz range [25]. Fig. 9 shows the active power drawn by the arc furnace. Note that the STATCOM is able to improve the line active power such that active power variations caused by the arc furnace do not propagate throughout the system as shown in Fig. 10. The simulation model and control scheme is described in detail in [28]. The dc capacitor voltages normally vary and are kept in relative balance through redundant state selection [20].

A. Dynamic Performance

To test the proposed fault detection and mitigation approach, a faulty switch was initiated at 2.5 s. Within 300 ms, the fault has been detected, the module removed, and the STATCOM restored to steady-state operation. This fault duration is longer than is necessary; the fault was intentionally left on to better illustrate its effect on the system and removal. The STATCOM bus voltage and line active powers are shown before the fault, during, and after the faulty module is removed (Figs. 11 and 12). Note that both the bus voltage and line active power are adversely affected during the fault. In both cases, the high frequency oscillations are increased. Once the faulty module is removed, the system returns to its pre-fault behavior. There is a small induced low-frequency oscillation that can be observed in the line active power, but this is rapidly damped by the STATCOM’s control.

The average dc link voltage before, during, and after the fault is shown in Fig. 13. During the fault, the dc voltage drops rapidly as the faulted module capacitor discharges. When the faulty module is removed, the average dc voltage drops to roughly 80% of the initial voltage, as expected. The continued variation in the dc link voltage is due to the continual variation of the arc furnace load that the STATCOM is compensating and is normal.

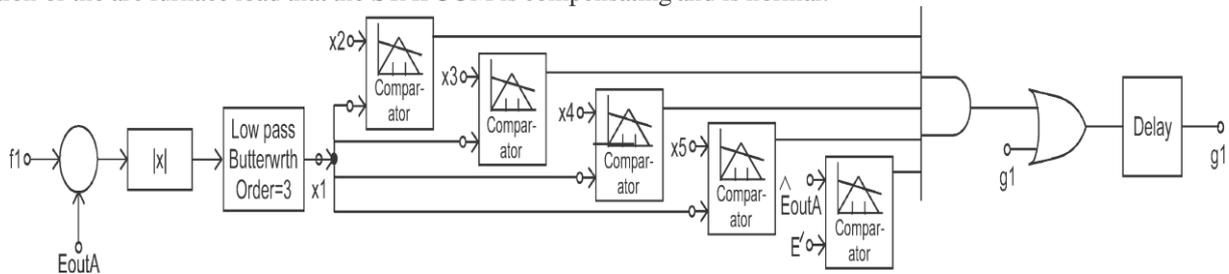


Fig. 7. Proposed fault detection and remediation control for cell 1.

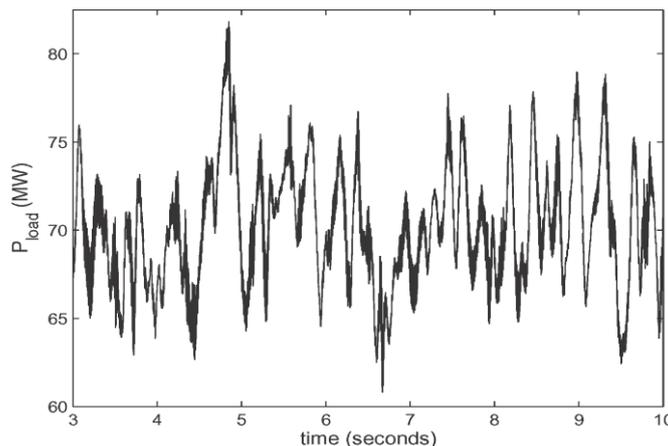


Fig. 8. Active power drawn by the arc furnace load.

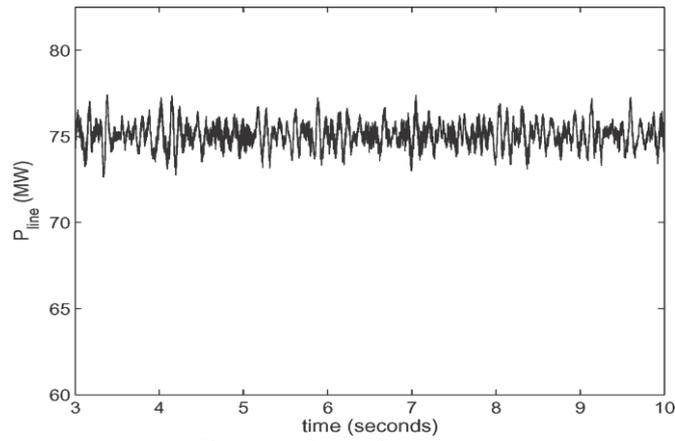


Fig. 9. Line active power.

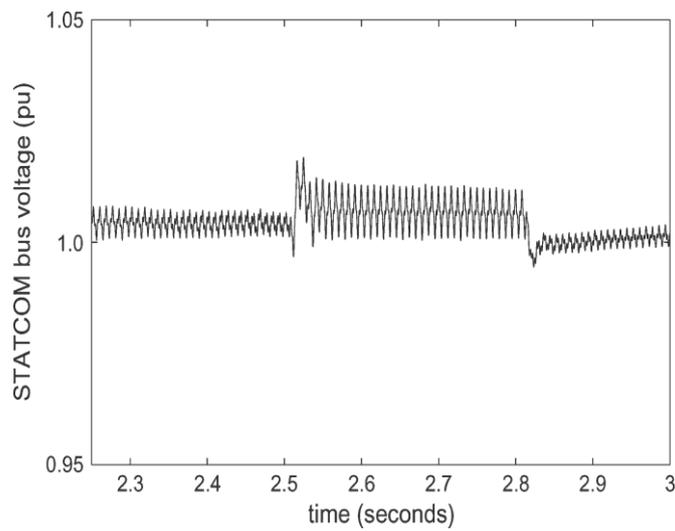


Fig. 10. STATCOM voltage before, during, and after fault.

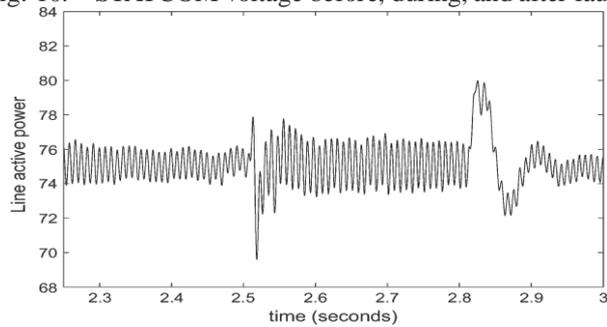


Fig. 11. Line active power before, during, and after fault.

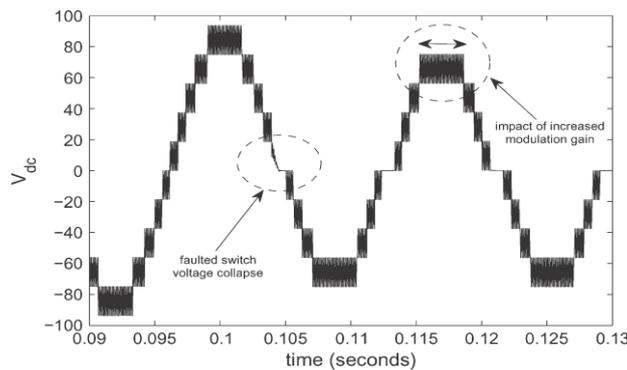


Fig. 12. Converter output with faulted cell.

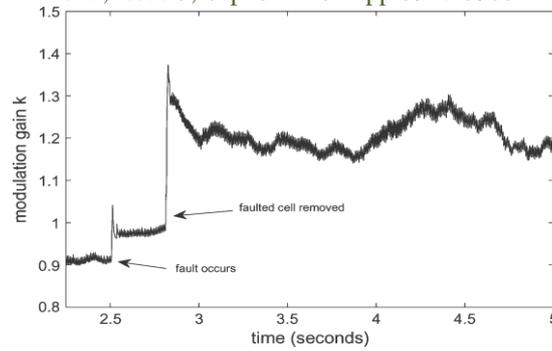
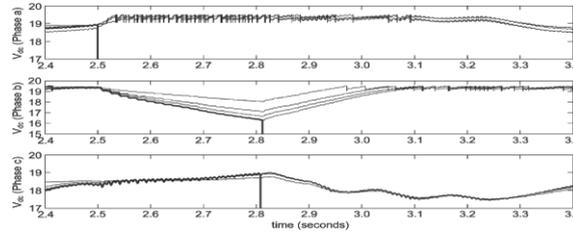
Fig. 13. Modulation gain k before, during, and after fault.

Fig. 14. Individual module capacitor voltages before, during, and after fault.

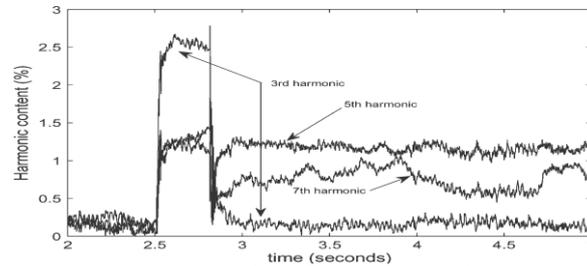


Fig. 15. Percent harmonic content of the faulty phase before, during, and after fault.

Fig. 14 shows two cycles of the STATCOM multilevel voltage output. There are several important aspects of this output waveform that have been highlighted. First, note the voltage collapse of the first level due to the faulted cell. This collapse in voltage will occur at the level that corresponds to the faulty cell. It is not possible to directly correlate the level number with the cell number (i.e., a collapse in level four does not necessarily indicate a fault in cell 4) because of the redundant state selection scheme that is used to balance the capacitor voltages.

A further aspect of note is the increase in length of the top-level duration. This is due to the increase in the modulation gain k due to the decrease in dc link voltage. Since the STATCOM output voltage is directly proportional to

$$V_{\text{stat}} = kV_{\text{dc}} \cos \alpha$$

Where k is the modulation gain and α is the phase angle. If V_{dc} decreases by 20%, then k must increase by 20% to compensate. An increase of this magnitude in modulation gain takes the PWM into over modulation where the magnitude of the reference waveform exceeds the magnitude of the carrier. This results in an increased length of time at higher voltage levels. Over modulation may also result in the increase of lower frequency harmonics. The modulation gain k is shown in Fig. 15. The individual module capacitor voltages in each phase for a faulty a phase switch are shown in Fig. 16. Note that the faulted module voltage decays rapidly at 2.5 s (when the fault was applied). The remaining capacitor voltages in phase a show significant “chopping” as the redundant state selection approach rapidly alternates between modules to maintain the average dc link voltage. A crowbar circuit is used with each module to limit the maximum dc voltage, leading to the chopping behavior. Phase b shows a continual decline in all of the capacitor voltages until the corresponding faulty module is removed at 2.8 s. The capacitor voltages increase until they are in the nominal range and then exhibit similar “chopping” until they are regulated. Phase c does not exhibit chopping because all of the individual cell voltages are of similar magnitude and do not exceed the crowbar maximum.

B. THD Performance

Harmonic injection is a concern with STATCOMs [29]. A harmonic analysis has been performed on the output voltage at the point of common coupling. One of the primary reasons for using a multilevel converter is the reduction in harmonic content in the output waveform. Fig. 17 shows the harmonic distortion levels at the STATCOM PCC before, during, and after the fault. Since this is measured at the PCC, the output waveform has already been filtered to remove high-frequency components. Before the fault, the THD level is less than 1%, which is quite good. During the fault, the THD increases to over 5%. When the fault is removed, the THD decreases and settles at approximately 2.5%, which is in the acceptable range for a 115-kV system [30]. Therefore, the loss of one of the cells does not necessitate the immediate removal of the STATCOM from service. The increase in THD after removing the faulty cell is due to several reasons. First, the

STATCOM filters were tuned to the resonant frequencies associated with the eleven-level converter and are not as effective when the converter topology changes to a 9-level. Second, the over modulation required for the 9-level converter increases the content of the lower frequency harmonics. While the third harmonic is quite high during the fault, it returns to pre-fault levels after the fault is cleared, whereas the fifth and seventh remain fairly high due to the over modulation. Even though they are increased over the pre-fault value, they still remain under the 1.5%-limit required of 115-kV systems [30].

V. EXPERIMENTAL RESULTS

To confirm the operation of the fault detection algorithm for cascaded H-bridge multilevel converters, an experimental prototype is constructed for applying and detecting different type of faults. The laboratory prototype of the STATCOM converter is shown in Fig. 18. The experimental rack consists of 36 Power ex CM75Du-24F IGBTs rated at 1200 V and 75 A for main switching devices. Passive components include a 1.2-mH, 45-A reactor and 18 electrolytic capacitors rated at 3900 μ F and 450 V. The IGBTs are driven by a Concept 6SD106E1 gate driver and controlled by a 320F2812 fixed-point digital signal processor (DSP).

For this prototype, three H-Bridge cells are cascaded to make a seven-level inverter for each phase of the STATCOM. During normal operation, the capacitor voltage for each cell is 30 V. A TMS320F2812 DSP is used for the calculation of the modulation, control, and fault detection algorithms. The H-bridges are constructed from the IGBTs switches and the gate signals of the IGBTs are delivered via fiber optic cables.

The output voltage of the converter during the normal operation, during the fault, and after removing the faulty cell is depicted in Fig. 19(a). A fault is applied to the second cell at point "F" as shown in the figure with the dashed line. Immediately after the fault is applied, the block capacitor begins to discharge. The capacitor voltage is shown in Fig. 19(b). Fig. 19(c) shows the output of the fault detection algorithm. When a fault is detected, g_2 (the gate signal of the by-pass circuit) becomes activated and it triggers the by-pass of the second cell and deactivates the PWM commands to the faulty H-bridge. Fig. 19(d) shows the filtered output of the STATCOM. Note that the amplitudes of the sinusoidal waveforms are nearly identical before and after fault detection and bypass. After detecting the fault and bypassing the faulty H-bridge, the modulation index is increased to compensate for the lost voltage levels in the output. In addition, the PWM switching patterns are modified based on existence of two cascaded H-bridges instead of three. This causes significant improvement in the output waveform of the converter.

Fig. 20 shows the same fault as in Fig. 19, except the fault bypass signal is intentionally delayed by several cycles to demonstrate the effect of changing the PWM pattern. Note that after the fault and discharge of the corresponding capacitor, the output waveform contains considerable distortion. However, modifying the PWM switching signals based on two cascaded H-bridges, the THD of the output waveform can be significantly decreased and the filtered output waveform become sinusoidal again.

VI. CONCLUSION

In this paper, a fault allocation strategy for a multilevel cascaded converter has been proposed. This approach requires no extra sensors and only one additional bypass switch per module per phase. The approach has been validated on a 115-kV system with a STATCOM compensating an electric arc furnace load. This application was chosen since the arc furnace provides a severe application with its non-sinusoidal, unbalanced, and randomly fluctuating load. The proposed approach was able to accurately identify and remove the faulted module. In addition, the STATCOM was able to re-main in service and continue to provide compensation without exceeding the total harmonic distortion allowances.

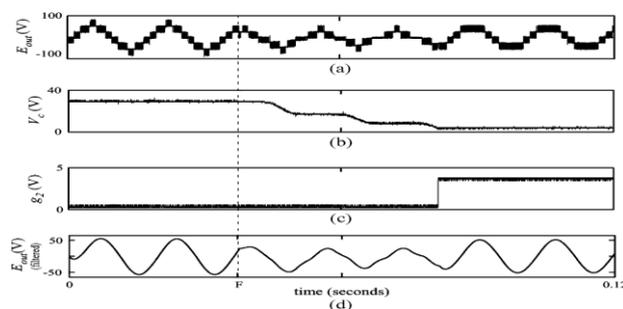


Fig. 16. Experimental STATCOM dynamics before, during, and after a fault is applied; (a) E_{out} , (b) capacitor voltage at faulted cell, (c) gating signal of cell 2 by-pass switch, and (d) E_{out} (filtered).

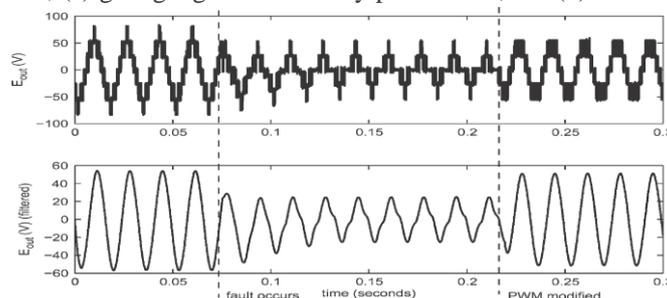


Fig. 17. Experimental STATCOM dynamics showing change in PWM

REFERENCES

- [1] C.-H. Liu and Y.-Y. Hsu, "Design of a self-tuning PI controller for a STATCOM using particle swarm optimization," *IEEE Trans. Ind. Elec- tron.*, vol. 57, no. 2, pp. 702–715, Feb. 2010.
- [2] D. Soto and T. C. Green, "A comparison of high-power converter topologies for the implementation of FACTS controllers," *IEEE Trans. Ind. Electron.*, vol. 49, no. 5, pp. 1072–1080, Oct. 2002.
- [3] J. A. Barrena, L. Marroyo, M. Á. Rodríguez Vidal, and J. R. Torrealday Apraiz, "Individual voltage balancing strategy for PWM cascaded H-bridge converter-based STATCOM," *IEEE Trans. Ind. Electron.*, vol. 55, no. 1, pp. 21–29, Jan. 2008.
- [4] Y. Liu, A. Q. Huang, W. Song, S. Bhattacharya, and G. Tan, "Small- signal model-based control strategy for balancing individual DC capacitor voltages in cascade multilevel inverter-based STATCOM," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2259–2269, Jun. 2009.
- [5] T. A. Meynard, M. Fadel, and N. Aouda, "Modeling of multilevel convert- ers," *IEEE Trans. Ind. Electron.*, vol. 44, no. 3, pp. 356–364, Jun. 1997.
- [6] C. Turpin, P. Baudesson, F. Richardeu, F. Forest, and T. A. Meynard, "Fault management of multicell converters," *IEEE Trans. Ind. Electron.*, vol. 49, no. 5, pp. 988–997, Oct. 2002.
- [7] S. Wei, B. Wu, F. Li, and X. Sun, "Control method for cascaded H-bridge multilevel inverter with faulty power cells," in *Proc. Appl. Power Electron. Conf. Expo.*, Feb. 2003, vol. 1, pp. 261–267.
- [8] S. Li and L. Xu, "Fault-tolerant operation of a 150 kW 3-level neutral point clamped PWM inverter in a flywheel energy storage system," in *Conf. Rec. 36th IEEE IAS Annu. Meeting*, Chicago, IL, Oct. 2001, pp. 585–588.
- [9] B. McGrath and D. Holmes, "Multicarrier PWM strategies for multilevel converters," *IEEE Trans. Ind. Electron.*, vol. 49, no. 4, pp. 858–867, Aug. 2002.
- [10] J. I. Leon, S. Vazquez, S. Kouro, L. G. Franquelo, J. M. Carrasco, and J. Rodriguez, "Unidimensional modulation technique for cascaded multi- level converters," *IEEE Trans. Ind. Electron.*, vol. 56, no. 8, pp. 2981–2986, Aug. 2009.
- [11] P. Lezana, J. Rodriguez, R. Aguilera, and C. Silva, "Fault detection on multi cell converter based on output voltage frequency analysis," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 2275–2283, Jun. 2009.
- [12] Task Force on Harmonics Modeling and Simulation, "Modeling devices with nonlinear voltage-current characteristics for harmonic studies," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1802–1811, Oct. 2004.
- [13] F. Richardeau, P. Baudesson, and T. Meynard, "Failure-tolerance And remedial strategies of a PWM multicell inverter," *IEEE Trans. Power Electron.*, vol. 17, no. 6, pp. 905–912, Nov. 2002.
- [14] S. Khomfoi and L. Tolbert, "Fault diagnosis and reconfiguration for multilevel inverter drive using AI-based techniques," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 2954–2968, Dec. 2007.
- [15] M. Ma, L. Hu, A. Chen, and X. He, "Reconfiguration of carrier-based modulation strategy for fault tolerant multilevel inverters," *IEEE Trans. Power Electron.*, vol. 22, no. 5, pp. 2050–2060, Sep. 2007.
- [16] S. Ceballos, J. Pou, E. Robles, I. Gabiola, J. Zaragoza, J. L. Villate, and D. Boroyevich, "Three-level converter topologies with switch breakdown fault-tolerance capability," *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 982–995, Mar. 2008.
- [17] P. Barriuso, J. Dixon, P. Flores, and L. Morán, "Fault-tolerant reconfig- uration system for asymmetric multilevel converters using bidirectional power switches," *IEEE Trans. Ind. Electron.*, vol. 56, no. 4, pp. 1300–1306, Apr. 2009.
- [18] P. Lezana and G. Ortiz, "Extended operation of cascade multicell con- verters under fault condition," *IEEE Trans. Ind. Electron.*, vol. 56, no. 7, pp. 2697–2703, Jul. 2009.
- [19] F. Z. Peng, J. S. Lai, W. McKeever, and J. VanCoevering, "A multilevel voltage source inverter with separate dc sources for static VAr generation," *IEEE Trans. Ind. Appl.*, vol. 32, no. 5, pp. 1130–1138, Sep. 1996.

About Authors



B. Bansiraj was born in 1985. He Received B.Tech (Electrical & Electronics Engineering) degree from Jawaharlal Nehru Technological University college of engineering, Kakinada, A.P., India in 2006. He has 4 years teaching experience in various engineering colleges in India. Presently He is Pursuing M.Tech (Power Electro-nics & Electrical Drives) from Mother Teresa Institute of Science & Technology, Sathupally, India affiliated by JNTU Hyderabad, India. His area of interest is Power Electronics, Power System.



Mr. M. LOKYA was born in 1984. He graduated from KAKATIYA UNIVERSITY, in the year 2005. He received M.Tech degree from Jawaharlal Nehru Technological University, Hyderabad in the year 2011. He is presently working as Assistant Professor in the Department of Electrical and Electronics Engineering at Mother Teresa Institute of Science and Technology, Andhra Pradesh, India.

An Electrical Model for Bronchitis and Emphysema of Human Respiratory System

¹Tinku Biswas, ²Surajit Bagchi, ³Swarup Sarkar, ⁴Akash Ku. Bhoi, ⁵Saikat Chatterjee

^{1,3,4,5} Department of AE&I Engg, Sikkim Manipal Institute of Technology (SMIT), Majitar

²Department of AE& I Engg, Heritage Institute of Technology, Kolkata

ABSTRACT: In this paper, we have developed an electrical model which mimics two respiratory diseases. The electrical circuit is analogous to the actual anatomical structure of the breathing system and composed of the passive electrical components (R, L and C). The transfer function is derived from the electrical model. Different time-domain responses are observed from the transfer function to study the alveoli and respiratory tract conditions for qualitative assessment of bronchitis (due to narrowed bronchial tubes) and emphysema (due to air pollutants, smoking etc.) For each of the diseases we have considered its three different qualitative stages or levels e.g. moderate, good and poor for productive analysis.

Keywords: Bronchitis, Emphysema, Transfer function, Time-domain responses.

I. INTRODUCTION:

The respiratory system (or ventilatory system) is the biological system of an organism that introduces respiratory gases to the interior and performs gas exchange. In humans and other mammals, the anatomical features of the respiratory system include airways, lungs, and the respiratory muscles. Molecules of oxygen and carbon dioxide are passively exchanged, by diffusion, between the gaseous external environment and the blood. This exchange process occurs in the alveolar region of the lungs. Other animals, such as insects, have respiratory systems with very simple anatomical features, and in amphibians even the skin plays a vital role in gas exchange. Plants also have respiratory systems but the directionality of gas exchange can be opposite to that in animals. [14]

The respiratory control system can also, under proper conditions, exhibit damped and sustained oscillation. Douglas and Haldane (1908) showed that after a period of voluntary hyperventilation several cycles of damped oscillation often occur [3], and in the clinical abnormality known as Cheyne-Stoke's respiration, the system continually overshoots and undershoots the regulated level and, thus, exhibits sustained oscillation. This study was designed to investigate the human respiratory control system. Its purpose has been, first, to derive the basic equations of the system and, secondly, to investigate the system as a biological regulator. This concept is not new; however the system has been analyzed mathematically only roughly. A simplified analysis was made by Grodins in 1954 in which CO₂ was the only controller of ventilation considered; the tissue elements were lumped into a single reservoir, blood flow was held constant, and circulation times were considered to be infinitely short. Horgan and Lange (1963)[4] added circulation times and oxygen control to this basic model in order to study periodic breathing. Defares et al. (1960) extended Grodins' model by dividing the tissue reservoir into two distinct compartments, brain and body tissues and considering cerebral blood flow as a function of arterial Pco₂. However, the effects of oxygen as a controller of ventilation and the effect of time delays in the transport of gases from the lungs to the two tissue reservoirs were not considered in this model because of their unimportance to the CO₂ inhalation studies with which this model was concerned. These factors have all proved to be very important in describing the respiratory control system. They will be considered in detail in this analysis.[2][1]

II. RESPIRATORY PROBLEMS

In our present work we are concentrating on two major conditions i) breathing problems due to narrowed bronchial tubes ii) emphysema due to air pollutants etc.[15] Breathing problems occur when the bronchial tubes become constricted, or when the alveoli lose their elasticity and have difficulty forcing the carbon dioxide out.[8][12]

Acute bronchitis is an inflammation of the bronchial tubes of the lungs. Severe bronchitis (usually caused by a viral infection) narrows these passages and secretes mucus. As a result, the impedance (R_b and L_b) of the electrical circuit, equivalent to the bronchial airways, rise abnormally as shown in figure 1. The increase in impedance causes a reduction in current flow through the circuit. [10]

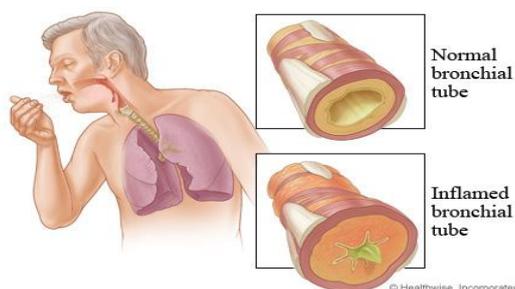


Fig.1: Normal and abnormal bronchial tubes [4]

We studied the status of alveoli due to extreme atmospheric pollution and human habits (like smoking etc.) Millions of tiny, flexible air sacs called alveoli connect to the bronchial tubes and blood vessels. This is where fresh air (oxygen) is exchanged with carbon dioxide from the blood. Air pollutants stimulate the production of elastase, which weakens the elasticity of the alveoli. When a substantial amount of elasticity is lost, the alveoli have difficulty pushing the carbon dioxide out, and one may have a dangerous condition called emphysema. The alveolar elasticity is described through its compliance (Fig.2) i.e. in the electrical circuit the capacitance (C_{bo}) decreases.[10][7]

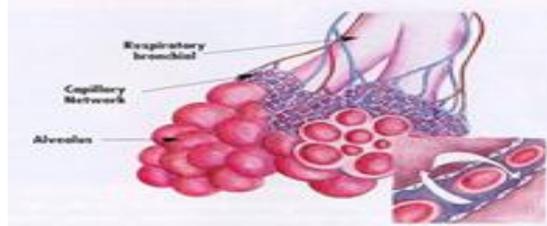


Fig.2: Status of alveoli

In first step, the equivalent electrical circuit of human respiratory system is designed considering the bronchial airways as the resistances and alveoli as the capacitors of the electrical circuit.

In second step, the transfer function of the equivalent electrical circuit is derived.

In third step, we studied the performance of the respiratory system through stability analysis.

III. THEORY

In the equivalent circuit, the electrical passive components represent different parts of lungs. R_t represents the trachea (an air way with largest cross-sectional area among the respiratory air ways). R_{b1} and R_{b2} are meant for the main bronchi (both left and right) (Bronchi are air paths with smaller cross-sectional areas relative to the trachea). [6][1]

The assembly, comprising of bronchiole and cluster of alveoli (Fig.2) is shown through Z_{bo} in Fig.3. Z_{bo} is a series combination of a resistor and an effective capacitor (So, alveolus, acting as an air sack, is supposed to carry the same quantity of air. The alveolus may be well thought-out an electrical analog of a capacitor.)

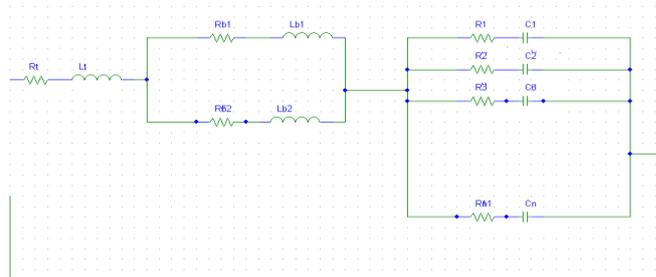


Fig.3: Electrical circuit of the respiratory system

IV. TRANSFER FUNCTION OF THE RESPIRATORY SYSTEM

We have derived the transfer function of the respiratory system based on the electrical circuit (Fig.2). It is known that alveolar number is closely related to the total lung volume. Hence one can say that larger the volume of lung more will be the alveoli. Here, we have considered the mean alveolar number as 480 million (range 274-790 million, coefficient of variation 37%). The mean volume of an alveolus is constant irrespective of the lung size [1, 2].[11]

$$z(s) = R_t + L_t s + \frac{\{R_{b1}R_{b2} + s(L_{b1} + L_{b2}) + L_{b1}L_{b2}s^2\}}{\{R_{b1} + R_{b2} + s(L_{b1} + L_{b2})\}} + (sRC_{bo} + 1)/nsC_{bo}$$

$$z(s) = \{s^3(nC_{bo}L_tL_{b1} + nC_{bo}L_tL_{b2} + nC_{bo}L_{b1}L_{b2}) + s^2(nC_{bo}L_tR_{b1} + nC_{bo}L_tR_{b2} + nC_{bo}R_tL_{b1} + nC_{bo}R_tL_{b2} + nC_{bo}L_{b2} + nC_{bo}L_{b1} + RC_{bo}L_{b1} + RC_{bo}L_{b2}) + s(nC_{bo}R_tR_{b1} + nC_{bo}R_tR_{b2} + nC_{bo}R_{b1}R_{b2} + RC_{bo}R_{b1} + RC_{bo}R_{b2} + L_{b1} + L_{b2}) + (R_{b1} + R_{b2})\} / \{nsC_{bo}(R_{b1} + R_{b2} + s(L_{b1} + L_{b2}))\}$$

$$y(s) = \frac{1}{z(s)}$$

$$y(s) = \{nsCbo(Rb1 + Rb2 + s(Lb1 + Lb2))\} / \{s^3(nCboLtLb1 + nCboLtLb2 + nCboLb1Lb2) + s^2(nCboLtRb1 + nCboLtRb2 + nCboRtLb1 + nCboRtLb2 + nCboLb2 + nCboLb1 + RCboLb1 + RCboLb2) + s(nCboRtRb1 + nCboRtRb2 + nCboRb1Rb2 + RCboRb1 + RCboRb2 + Lb1 + Lb2) + (Rb1 + Rb2)\}$$

The transfer function = $\frac{I(s)}{V(s)} = Y(s)$

SIMULATION TOOLS:

The stability analysis of the equivalent electrical circuit is done in MatLab platform. Here, we used MatLab7.0 software: To study the response of emphysema and To study the response of bronchitis for step and ramp input.

PARAMETERS:

- Impedance of trachea – Rt and Lt
- Impedance of main bronchi (left) - Rb1 and Lb1
- Impedance of main bronchi (right) -Rb2 and Lb2
- Resistance of the bronchial airways –R
- Compliance of the alveoli- Cbo

Rt (unit)	Rb1 (unit)	Rb2 (unit)	R (unit)	Lt (unit)	Lb1 (unit)	Lb2 (unit)	Cbo (unit)
4.5	9	9	0.5	0.043	0.17	0.17	0.01

V. RESULTS AND DISCUSSIONS

Condition 1: To study the response of emphysema

We have observed the responses of the respiratory system for emphysema with three different stages. i) average ii) good and iii) poor alveoli status. For average status we have assumed that the air ways impedances are constant and the electrical compliance is averaged with c=0.001 unit. Figs.4 (a) and (b) give the ramp and step responses. Output due to the ramp input quickly attains the steady state whereas the step sensation gives a damped sinusoid output resembling a second order underdamped system.

Stage1: We studied the performance of the respiratory system through stability analysis taking values of Rt= 4.5unit, Rb1=9 unit, Rb2=9 unit, R=0.5 unit, Lt=0.043 unit, Lb1=0.17 unit, Lb2=0.17 unit, Cbo= 0.001 unit in the electrical circuit [Fig 4] as standard and we got transfer function as

$$\frac{0.00034s^2 + 0.018s}{0.00004352s^3 + 0.002814s^2 + 0.511s + 18}$$

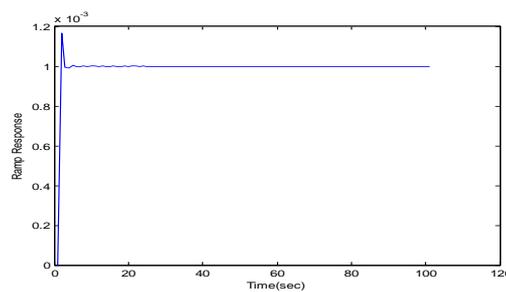


Fig.4 (a)

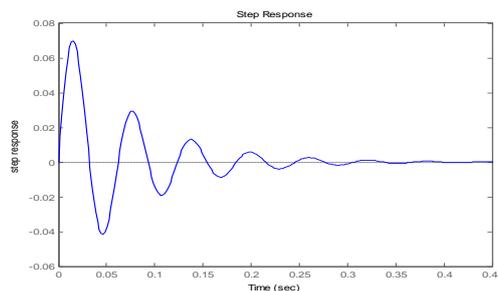


Fig. 4 (b)

Fig. 4 Average alveoli condition with compliance C=0.001 unit (a) Ramp response (b) Step response.

Stage2: Ramp and step responses are exact replica of each other when the patient's alveoli are at good condition as reflected in figs5 (a) and (b). Here keeping the air tract impedances unchanged we have altered the alveolar compliance to 0.01 unit. The transfer function was derived as

$$\frac{0.0034s^2 + 0.18s}{0.0004352s^3 + 0.02814s^2 + 2.05s + 18}$$

taking Cbo=0.01 unit. The transients are dying out very quickly indicating a more stable alveolar state.

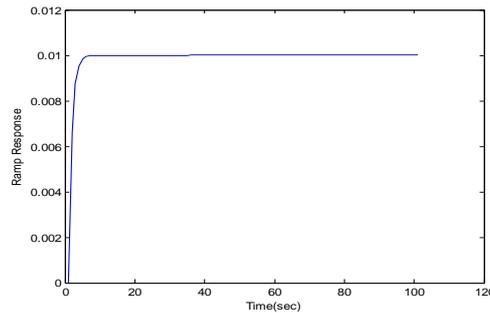


Fig.5 (a)

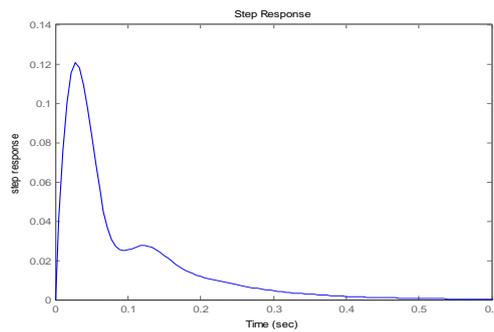


Fig. 5 (b)

Fig. 5 Average alveoli condition with compliance C=0.01 unit (a) Ramp response (b) Step response.

Stage3: Alveoli condition deteriorates due to air pollutants, as a result the alveolar compliance decreases. We have followed the same analysis with ramp and step inputs choosing a value of C_{bo}=0.0001 unit.

Here we were taking Cbo=0.0001 in the transfer function (became

$$\frac{0.000034s^2 + 0.018s}{0.00004352s^3 + 0.0002814s^2 + 0.2571s + 18}$$

The ramp and step responses are shown in Fig. 6 (a) and (b).

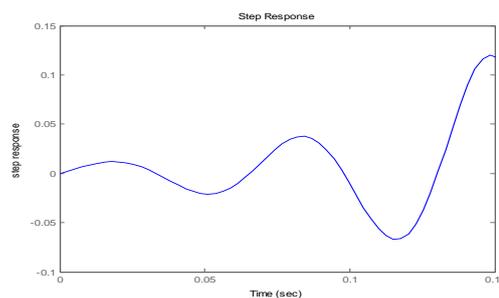
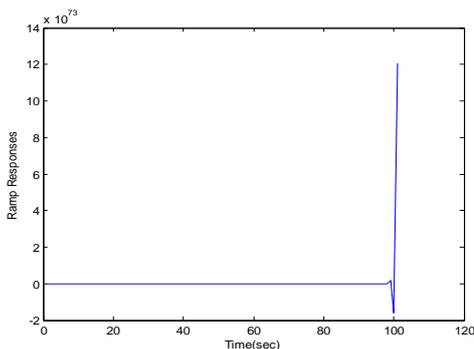


Fig. 6(a) & (b)

Fig.6 Depiction of poor alveoli condition with $c= 0.0001$ unit through (a) ramp response (b) step response

Condition 2: To study the response of bronchitis

To study the responses of bronchitis, we changed the values of R, Rt, Rb1, Rb2, Lt, Lb1, Lb2 to 1000 times to the standard values of each one respectively in the transfer fuction taking $C_{bo}=0.001$ unit. The required tranfer function to plot step response [Fig 7] and ramp response[fig.8]becomes

$$\frac{0.34s^2 + 18s}{43.52s^3 + 2474.34s^2 + 171340s + 18000}$$

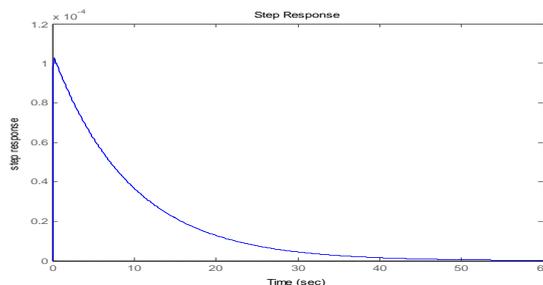


Fig. 7 Step responses with $c = 0.001$ unit

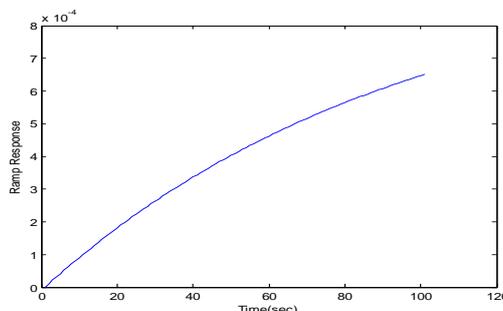


Fig 8: Ramp response taking all values of R and L 1000times to standard and $C_{bo}=0.001$ unit.

In the Fig. 8 the transients are decreasing towards time axis very slowly. It meets time axis around 50 sec, which is very slow. The ramp response also deviating from the standard one. The impedance (R_b and L_b) of the electrical circuit, was changed, narrows the passages and secretes mucus of the lung.

VI. CONCLUSION

This paper discussed about the study of i) breathing problems due to narrowed bronchial tubes ii) emphysema due to air pollutants etc. With increasing the impedances the system attains the stability very slowly [Fig. 7 and 8]. It takes more time (approximate 49.55 seconds more) to stable. After discussing Fig. 4, 5 and 6 we can conclude that the alveolar compliance is directly proportional with the stability of the respiratory system. The transfer function can be used to study the *chronic obstructive pulmonary disease (COPD)*, Pneumonia.

I. REFERENCES

- [1] DEARUS, J. G., DERKSON, H. E., and DUYFF, J. W., 1960, Acta Physiol. et Pharmacol. Neerl.9, 327.
- [2] GRODINS, F. S., GRAY, J. S., SCHROEDER, K. R., NoRINs, A. L., and JONES, R. W., 1954, 1. Appl.Physiol. 7, 283.
- [3] DOUGLAS, C. G., and HALDANE, J. S., 1908, J. Physiol., 38, 401.
- [4] HORGAN, J. D., and LANGE, R. L., 1963, Institute of Electrical and Electronic Engineers Convention Record, pt. 9, 149.
- [5] Matthias Ochs, Jens R. Nyengaard, ,Anja Jung, Lars Knudsen, Marion Voigt, Thorsten Wahlers, Joachim Richter and Hans Jørgen G. Gundersen, *The Number of Alveoli in the Human Lung* , *American Journal of Respiratory and Critical Care Medicine*
- [6] Ochs M, Nyengaard JR, Jung A, Knudsen L, Voigt M, Wahlers T, Richter J, Gundersen HJ. *The number of alveoli in the human lung*, *Department of Anatomy, Division of Electron Microscopy, University of Göttingen, Göttingen, Germany*. [7] <http://lung-conditions.factoidz.com/asthma-vs- bronchitis/>
- [8] <http://e-pena.blogspot.in/2011/03/viral-bronchitis.html>
- [9] <http://kats-korner.com/health/health.html>
- [10] The Human Respiratory System. users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Pulmonary.html

- [11] A highly attenuated recombinant human respiratory syncytial virus lacking the G protein induces long-lasting protection in cotton rats. Myra N Widjoatmodj, Jolande Boes, Marleen van Bers, Yvonne van Remmerden, Paul JM Roholl **and** Willem Luytjes. *Virology Journal* 2010, 7:114 doi: 10.1186/1743-422X-7-114
- [12] Occupational airway diseases in the metalworking industry Part 1: Respiratory infections, pneumonia, chronic bronchitis and emphysema Original Research Article *Tribology International, Volume 18, Issue 3, June 1985, Pages 169-176* E.O. Bennett, D.L. Bennett
- [13] Sulfated oligosaccharides isolated from the respiratory mucins of a secret or patient suffering from chronic bronchitis Original Research Article *Biochimie, Volume 85, Issues 3-4, March-April 2003, Pages 369-379* Sophie Degroote, Emmanuel Maes, Pascale Humbert, Philippe Delmotte, Geneviève Lamblin, Philippe Rouse
- [14] Management of chronic bronchitis and acute exacerbations of chronic bronchitis Original Research Article *International Journal of Antimicrobial Agents, Volume 9, Issue 2, September 1997, Pages 83-93* Meyer Balter, Ronald F. Grossman
- [15] Antibiotic prescribing in ambulatory care settings for adults with colds, upper respiratory tract infections, and bronchitis Original Research Article *Clinical Therapeutics, Volume 24, Issue 1, January 2002, Pages 170-182* Ron Cantrell, A.Fred Young, Bradley C. Martin

An Individual Channel Designing For Integrated Multi-Port Dc-Dc Converter for Renewable Energy Applications

Devanaboyina S Singh¹, Vargil Kumar², D Srinivasa Rao³

¹(PG Scholar, Department of EEE, Gudlavaleru college of Engg / Jntuk, gudlavaleru, AP, India)

²(Assoc Prof, Department of EEE, Gudlavaleru college of Engg/ Jntuk, India)

ABSTRACT: This paper proposes a novel converter topology that Interfaces four power ports: two sources, one bidirectional storage port, and one isolated load port. The proposed four-port dc/dc converter is derived by simply adding two switches and two diodes to the traditional half-bridge topology. Zero-voltage switching is realized for all four main switches. Three of the four ports can be tightly regulated by adjusting their independent duty-cycle values, while the fourth port is left unregulated to maintain the power balance for the system. In this a new control method individual channel designing is implemented to get efficient results, the simulation results are shown below.

Keywords: DC-DC converter, half-bridge, multiple-input Single-output (MISO), multiport, zero-voltage switching (ZVS).

I. INTRODUCTION

As interest in renewable energy systems with various sources becomes greater than before, there is a supreme need for integrated power converters that are capable of interfacing, and concurrently, controlling several power terminals with low cost and compact structure. Meanwhile, due to the intermittent nature of renewable sources, a battery backup is normally required when the ac mains is not available. This paper proposes a new four-port-integrated dc/dc topology which is suitable for various renewable energy harvesting applications. An application interfacing hybrid photovoltaic (PV) and wind sources, one bidirectional battery port, and an isolated output port is given as a design example. It can achieve maximum power-point tracking (MPPT) for both PV and wind power simultaneously or individually, while maintaining a regulated output voltage. Compared to the effort spent on the traditional two-port converter, less work has been done on the multiport converter. But, due to the advantages like low cost and compact structure, multiport converters are reported to be designed for various applications, such as achieving three bus voltages of 14 V/42 V/H.V. (high voltage of around 500 V) in electric vehicles or hybrid electric vehicles interfacing the PV panel systems PV energy harvesting with ac mains or the battery backup hybrid fuel cell and battery systems, and hybrid ultra capacitor and battery systems. From the topology point of view, multi input converters based on buck, boost, and buck-boost topologies have been reported in. The main limitation of these configurations is the lack of a bidirectional port to interface storage device.

Multiport converters are also constructed out of a multi winding transformer based on half-bridge or full

bridge topologies. They can meet isolation requirement and also have bidirectional capabilities. However, the major problem is that they use too many active switches, in addition to the bulky transformer, which cannot justify the unique features of low component count and compact structure for the integrated multiport converter. The proposed four-port dc/dc converter has bidirectional capability and also has one isolated output. Its main components are only four main switches, two diodes, one transformer, and one inductor. Moreover, zero-voltage switching (ZVS) can be achieved for all main switches to allow higher efficiency at higher switching frequency, which will lead to more compact design of this multiport converter. The control design is also investigated based on the modeling of this modified half-bridge topology. In addition, a decoupling network is introduced to allow the separate controller design for each power port. Finally, a prototype has been built to verify the four-port converter's circuit operation and control capability. The proposed converter is a valuable candidate for low-power renewable energy harvesting applications.

The toolbox is inspired on a new approach for multivariable control systems, referred as Individual Channel Design (ICD). ICD is a novel analytical framework that allows the analysis and synthesis of multivariable control systems under the context of the Multivariable Structure Function (MSF) by applying classical techniques based on the Bode and Nyquist plots. With the help of this framework it is possible to investigate the potential and limitations for feedback design of any multivariable linear time-invariant control system. Although ICD is in principle a feedback structure based on diagonal controllers, it can be applied to any cross coupled multivariable system. It is based on the definition of individual transmission channels. In this context the control design is an interactive process that involves the required specifications, plant characteristics, and the multivariable feedback design process itself. Once the channels are defined it is possible to form a feedback loop with the compensator specially designed to meet *customer* specifications. In this manner the multivariable control design problem is reduced to the design of a single-input single-output control for each channel [4,5]. ICD has been reported in some control strategies, such as in small scale power networks with embedded generation [2], in the automotive and the aerospace industry [6]. So far, this toolbox has been used in different control tasks, from induction motors [9,10], synchronous generators [11], to submarines [7].

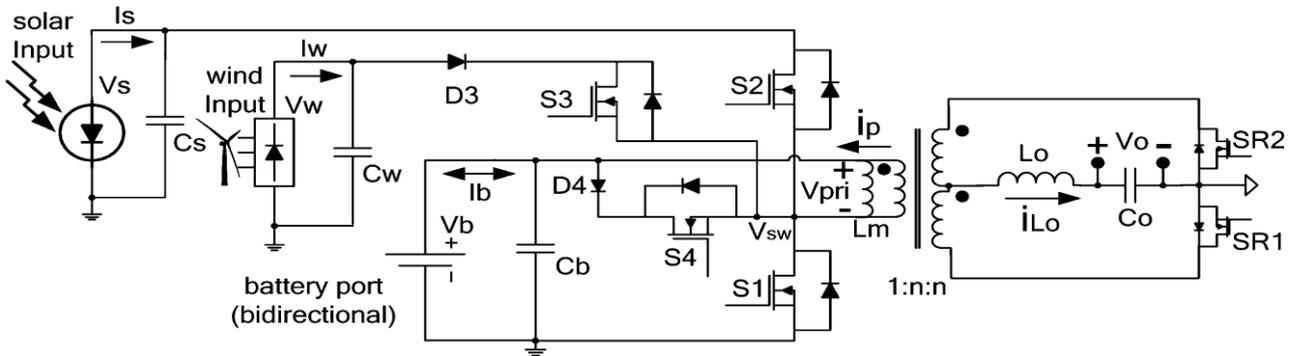


Fig. 1. Four-port half-bridge converter topology, which can achieve ZVS for all four main switches (S_1 , S_2 , S_3 , and S_4) and adopts synchronous rectification

for the secondary side to minimize conduction loss.

II. Topology And Circuit Analysis

The four-port topology is derived based on the traditional two-port half-bridge converter, which consists of two main switches S_1 and S_2 . As shown in Fig. 1, one more input power port can be obtained by adding a diode D_3 and an active switch S_3 . Another bidirectional power path can be formed by adding a freewheeling branch across the transformer primary side, consisting of a diode D_4 and an active switch S_4 . As a result, the topology ends up with four active switches and two diodes, plus the transformer and the rectification circuit. The proposed converter topology is suitable for a number of power-harvesting applications, and this paper will target the hybrid PV wind application. It should be noted that since the wind turbine normally generates a three phase ac power, an ac/dc rectifier needs to be installed before this four-port dc/dc interface and after the wind turbine output. And then rectification stage can utilize either active power factor correction (PFC) or passive PFC. However, the ac/dc solution is beyond the scope of this paper.

2.1. Driving Scheme

Fig. 2 illustrates a possible modulation approach to realize the constant frequency pulse width modulation (PWM) control, Where V_{saw} is the sawtooth carrier waveform for modulation, V_{c1} , V_{c2} , and V_{c3} are control voltages derived from the voltage or current feedback controllers. By modulating these control voltages, driving signals for S_1 , S_2 , and S_3 can be generated, respectively. Then, by reversing S_1 and S_3 driving signals, S_4 and two SR signals can be obtained. It should be noted that S_2 , S_3 , and S_4 do not need to be gated ON at the same time; instead, S_3 is only required to turn ON a little earlier before S_2 turns OFF, and S_4 is only required to turn ON a little earlier before S_3 turns OFF. No dead time is necessary between S_2 and S_3 , nor between S_3 and S_4 ,

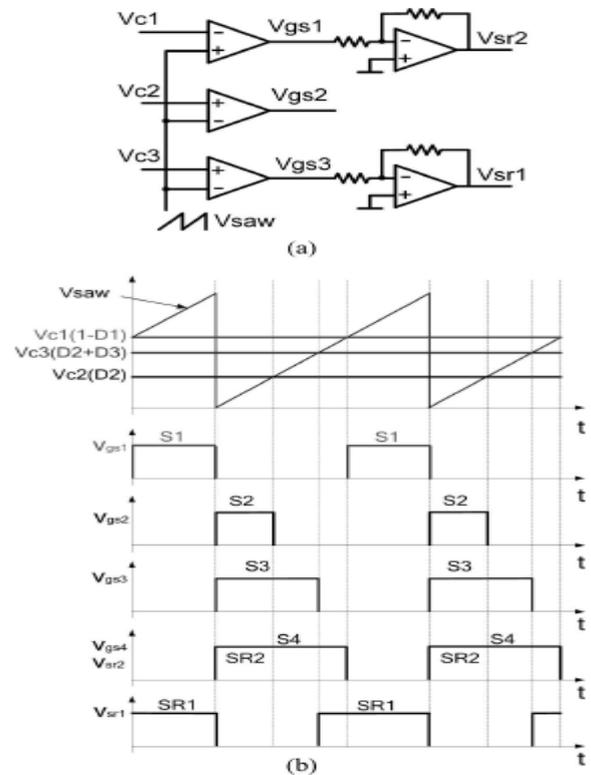


Fig. 2. Proposed modulation scheme. (a) PWM modulation circuits. (b) Driving signal key waveforms.

because the existence of diodes can prevent shoot-through problems. But the dead time between S_1 and S_2 and between S_1 and S_4 is necessary to prevent shoot-through, and also to create ZVS conditions for S_1 and S_2 .

2.2 Principle of Circuit Operation

The steady-state waveforms of the four-port converter are shown in Fig. 3, and the various operation stages in one switching cycle are shown in Fig. 4. To simplify the analysis of operation, components are considered ideal, except otherwise indicated. The main operation stages are described as follows.

Stage 1 (t0–t1): Before this stage begins, the body diode of

S1 is forced on to recycle the energy stored in the transformer leakage inductor, and the output is freewheeling. At time t0, S1 is gated ON with ZVS, and then, the leakage inductor is reset to zero and reverse-charged.

Stage 2 (t1–t2): At time t1, the transformer primary current increases to the reflected current of i_{Lo} , the body diode of SR2 becomes blocked, and the converter starts to deliver power to output.

Stage 3 (t2–t3): At time t2, S1 is gated OFF, causing the leakage current i_p to charge the S1 parasitic capacitor and discharge the S2, S3, and S4 parasitic capacitors.

Stage 4 (t3–t4): At time t3, the voltage across the S2 parasitic capacitor is discharged to zero, and the S2 body diode conducts to carry the current, which provides the ZVS condition for S2. During this interval, the output is freewheeling through SR1 and SR2 body diodes.

Stage 5 (t4–t5): At time t4, S2 is gated ON with ZVS, and

then, the leakage inductor is reset to zero and reverse-charged.

The output inductor current drop from t2 to t5 is due to the leakage inductor discharge/charge.

Stage 6 (t5–t6): At time t5, the transformer primary current increases to the reflected current of i_{Lo} , the body diode of SR1 is blocked, and the converter starts to deliver power to output.

Stage 7 (t6–t7): At time t6, S2 is gated OFF, thus causing the leakage current i_p to charge the S2 parasitic capacitor and discharge the S1 and D3 parasitic capacitors.

Stage 8 (t7–t8): At time t7, the voltage across D3 is discharged to zero, and then, D3 conducts. S3 is gated ON before this time; therefore, S3 has natural ZVS. Output inductor current freewheels through SR2 during this period.

Stage 9 (t8–t9): At time t8, S3 is gated OFF, thus causing the leakage current i_p to charge S2 and S3 parasitic capacitors and discharge S1 and D4 parasitic capacitors.

Stage 10 (t9–t10): At time t9, the voltage across D4 is discharged to zero and D4 conducts. Since S4 is gated ON before this time, the leakage current freewheels through D4 and S4, so that the leakage energy is trapped. On the secondary side, output inductor current freewheels through SR1 and SR2.

Stage 11 (t10–t11): At time t10, S4 is gated OFF, causing the trapped leakage energy to discharge the S1 parasitic capacitor and charge the S2, S3 and S4 parasitic capacitors.

Stage 12 (t11–t12): At time t11, the voltage across S1 is discharged to zero, and the S1 body diode conducts to carry the current, which provides ZVS condition for S1. During this interval, the output is freewheeling. This is the end of the switching cycle.

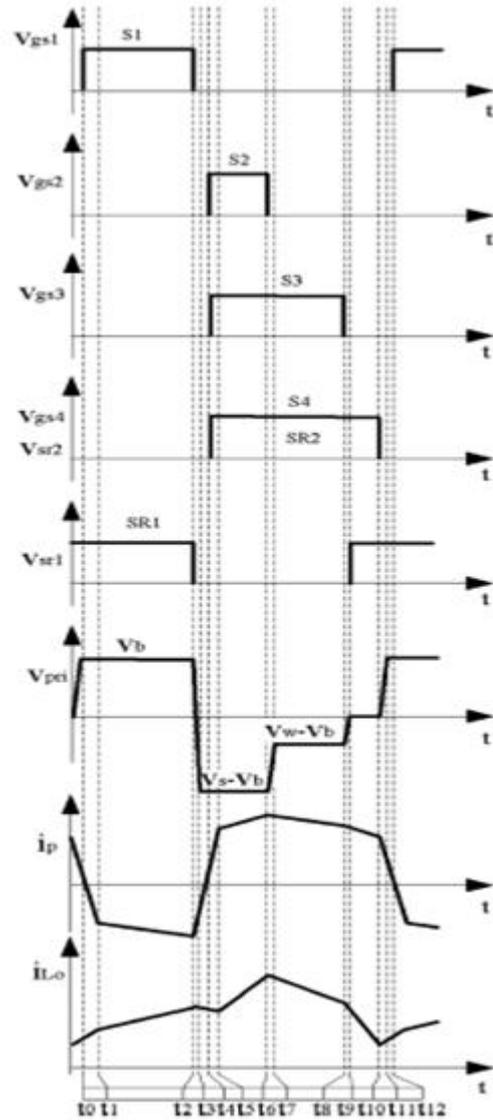


Fig. 3. Steady-state waveforms of the four-port half-bridge converter.

2.3 Steady-State Analysis

Assuming an ideal converter, the steady-state voltage governing relations between different port voltages can be determined by equating the voltage-second product across the converter's two main inductors to zero. First, using volt-second balance across the primary transformer magnetizing inductance LM in CCM, we have

$$V_b D_1 = (V_o - V_b) D_2 + (V_o - V_b) D_3 \quad (1)$$

Assuming CCM operation, the voltage-second balance across the load filter inductor Lo then yields

$$V_b D_1 + (V_s - V_b) D_2 + (V_o - V_b) D_3 = \frac{V_o}{n} \quad (2)$$

where n is the turns ratio of the transformer, V_s , V_w , V_b , V_o

are the solar input, wind input, battery, and output voltages,

respectively. The following equation is based on the power balance principle, by assuming a lossless converter, steady-state port currents can be related as follows

$$V_s I_s + V_\omega I_\omega = V_b I_b + V_o I_o \quad (3)$$

where I_s , I_w , I_b , I_o are the average solar input, wind input, battery bidirectional, and load currents, respectively. The battery current I_b is positive during charging and negative during discharging.

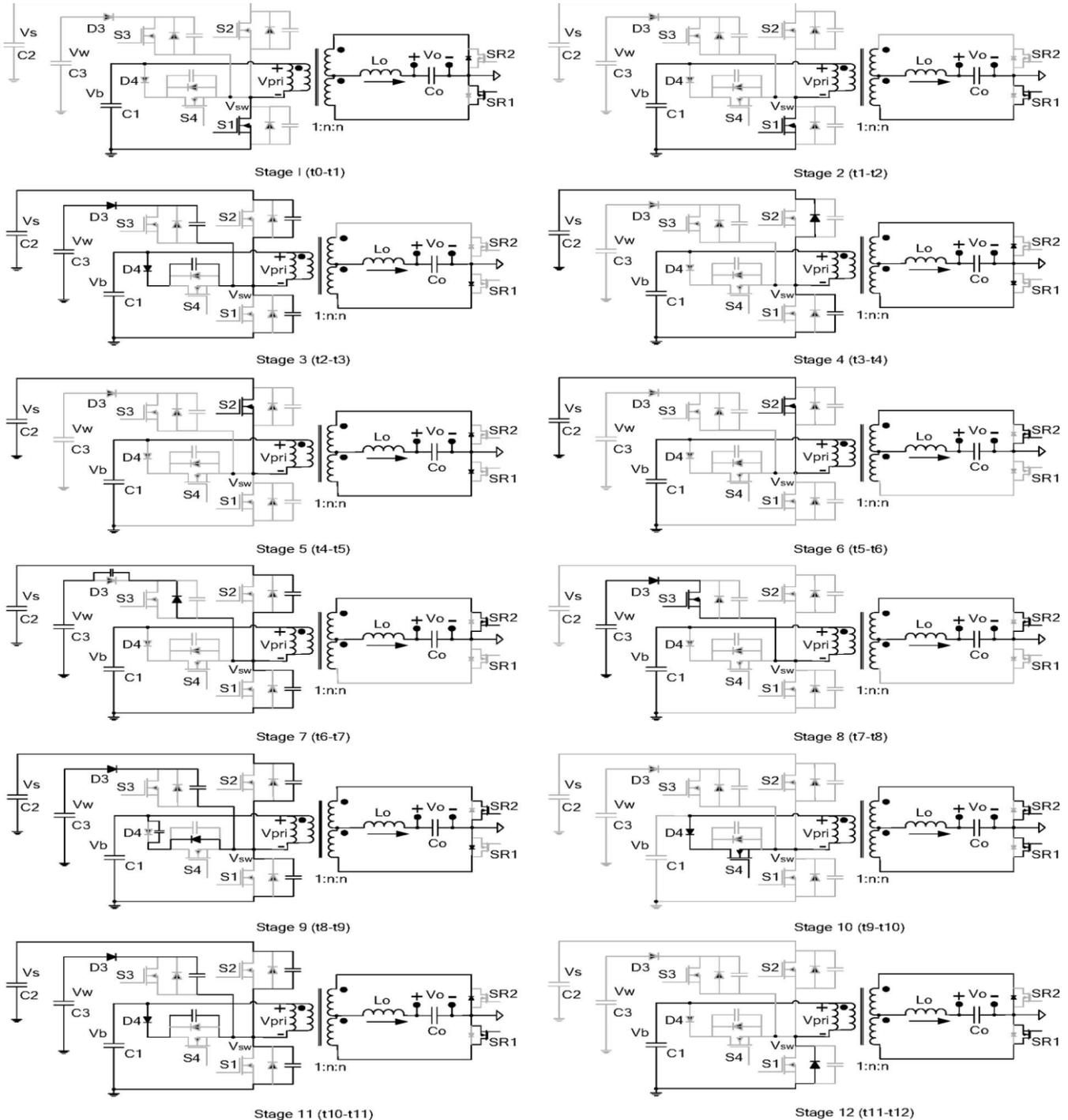


Fig. 4. Operation stages of the four-port half-bridge converter

2.4 ZVS Analysis

ZVS of the switches $S1$ and $S2$ can be realized through the energy stored in the transformer leakage inductor, while ZVS of $S3$ and $S4$ is always maintained, because the proposed will be forced on before the two switches turn ON After $S4$ is turned OFF, the leakage energy is released to discharge the $S1$ parasitic capacitor and charge $S2$, $S3$, and $S4$'s parasitic capacitors, to create the ZVS condition of $S1$. And the following condition should be

satisfied as C_{oss} , and I_M is the average transformer magnetizing current, which satisfies(4)

$$\frac{1}{2} L_k (I_M + nI_o)^2 > 2C_{oss} V_b^2 + C_{oss} V_b V_s + C_{oss} V_b V_\omega$$

$$I_M + nI_o > 0 \quad (4)$$

where Lk is the transformer leakage inductance, MOSFET parasitic capacitances of $S1$, $S2$, $S3$ and $S4$ are assumed to be equal as C_{oss} , and IM is the average transformer magnetizing current, which satisfies:

$$I_b = D_1(I_M + nI_o) + D_2(I_M + nI_o) + D_3(I_M + nI_o) \quad (5)$$

Rearranging (5), we can obtain IM as follows:

$$I_M = \frac{I_b + (D_1 - D_2 - D_3)nI_o}{(D_1 + D_2 + D_3)} \quad (6)$$

After $S1$ is turned OFF, the leakage energy will charge the $S1$ parasitic capacitor and discharge $S2$, $S3$, and $S4$'s parasitic capacitors to achieve ZVS for S

According to (7), when the load current I_o is small and the transformer magnetizing current IM is large, $IM - nI_o < 0$ cannot be met. In other words, ZVS of $S2$ will be lost. However, in most load/source conditions, ZVS of $S2$ is achievable. It should be noted that ZVS of $S3$ and $S4$ can be naturally achieved if the voltage relation $V_b < V_w < V_s$ is satisfied to ensure that the paralleling diodes will always be forced on before these switches turn ON. On one hand, $V_w < V_s$ is not difficult to meet since the solar port and wind port can be reversed if the wind port voltage V_w is larger than the solar port voltage V_s . Even if V_w is not always lower than V_s in the whole voltage ranges, the converter itself still works, but may lose some conduction period for the $S2$ branch depending on the driving overlap of $S2$ and $S3$. The solution is to change the driving scheme to avoid the $S2$ and $S3$ overlap. On the other hand, it is a step-down conversion from PV or wind port to battery port; therefore, the battery voltage V_b will be always lower than the PV voltage V_s and the wind source voltage V_w .

2.5. Circuit Design Considerations

When considering the semiconductor stresses, this modified half-bridge topology shows striking similarity to its traditional half-bridge counterpart. The major difference is that the transformer design of this four-port converter needs to allow for a dc current flow, and therefore, becomes similar to an inductor or a flyback transformer design. The dc biasing current rating is dictated by (6), which determines the amount of the air gap to be inserted. Other than the transformer, the circuit design and optimization technique used for the traditional half-bridge topology can be used here for this four-port topology, which provides great convenience for the practicing engineers to implement the power stage design.

III. CONTROL STRUCTURE AND DYNAMIC MODELING

The proposed converter has three freedoms to control the power flow of three power ports, while the fourth port is to maintain the power balance. That means the operating point of up to three ports can be tightly regulated, while the fourth port should be left "flexible" and would operate at any point that satisfies the power balance constraints. The choice of the flexible power port dictates the feedback control layout, which is based on different control objectives. For instance, if the battery is chosen to be left "flexible," the maximal power from the solar and wind sources can be tracked by their port voltages or currents independently, and the load voltage can be regulated by a voltage feedback as well.

3.1 Control Structure

Fig. 5 shows the control structure for the hybrid PV wind system. To the three feedback controllers' idc controller are given. Three feedback controllers are as follows: a solar voltage regulator (SVR), a wind voltage regulator (WVR), and an output voltage regulator (OVR). The OVR loop is simply a voltage-feedback loop, closed around the load port, and duty cycle $d1$ is used as its control input. The SVR loop is used to regulate the PV panel voltage to its reference value, which is provided by an MPPT controller. And the reference value represents an estimate of the optimal operating PV voltage; duty cycle $d2$ is used as its control input. The WVR loop is taking a very similar structure to SVR, except that its voltage reference represents the optimal operating voltage of the rectified wind turbine output voltage. The WVR loop is made to control $d3$. This control strategy allows $\frac{dx(t)}{dt} = Ax(t) + Bu(t)$, $y(t) = Ix(t)$ tightly regulated while maximizing the PV and wind power harvesting. In this system, the battery storage plays the significant role of balancing the system energy by injecting power at heavy loads and absorbing excess power when available PV and wind power exceeds the load demand.

3.2 Dynamic Modeling:

In order to design the SVR, WVR, and OVR controllers, a small signal model of the four-port converter is desired. The detailed modeling procedure can refer to [20], which is proposed for a three-port converter. And for this four-port converter, the general modeling procedure is very similar to [20]. Therefore, to avoid unnecessary repetition, only a brief introduction is given here. First, state-space equations for five energy storage elements during the four main circuit stages are developed. For the aforementioned mode of operation, these include the solar side capacitor C_s , the wind-side capacitor C_w , the transformer magnetizing inductor LM , the output inductor Lo , and the output capacitor Co . In the next step, state-space equations in the four main circuit stages (corresponding to the turn ON of four main switches) will be averaged, and then applied with the small signal perturbation. Finally, the first-order small-signal perturbation components will be collected to form the matrices A and B , which actually represent the converter power stage model. It should be noted that the symbolic derivation of these transfer functions is fairly tedious. Alternatively, the dynamics of the plant can be calculated by

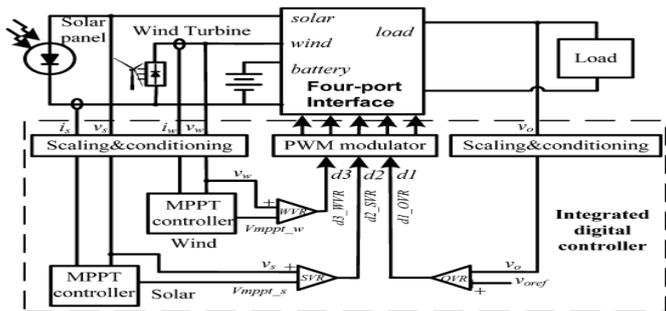


Fig. 5. Possible control structure to achieve MPPT for the PV panel and the wind turbine, meanwhile maintaining output voltage regulation. OVR, SVR, and WVR loops are to control d_1 , d_2 , and d_3 , respectively.

computer software like MATLAB. The resultant state-space averaging model takes the following form : (8) A

$$\begin{bmatrix} \frac{-1}{R_s C_s} & 0 & \frac{-D_2}{C_s} & \frac{-nD_2}{C_s} & 0 \\ 0 & \frac{1}{R_w C_w} & \frac{-D_3}{C_w} & \frac{-nD_3}{C_w} & 0 \\ \frac{D_2}{L_M} & \frac{D_3}{L_M} & 0 & 0 & 0 \\ \frac{nD_2}{L_o} & \frac{nD_3}{L_o} & 0 & 0 & \frac{-1}{L_o} \\ 0 & 0 & 0 & \frac{1}{C_o} & \frac{-1}{RC_o} \end{bmatrix}$$

$$\begin{bmatrix} 0 & \frac{-I_{L_m} - (nV_o / R)}{C_o} & 0 \\ 0 & 0 & \frac{-I_{L_m} - (nV_o / R)}{C_w} \\ \frac{-V_b}{L_m} & \frac{V_a - V_b}{L_M} & \frac{V_w - V_b}{L_M} \\ \frac{nV_b}{L_o} & \frac{n(V_a - V_b)}{L_o} & \frac{n(V_w - V_b)}{L_o} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\hat{u} = \begin{bmatrix} \hat{d}_1(t) \\ \hat{d}_2(t) \\ \hat{d}_3(t) \end{bmatrix}$$

where $\hat{x}(t)$ is a matrix containing the small signal state variables $\hat{v}_s(t)$, $\hat{v}_w(t)$, $\hat{i}_{Lm}(t)$, and $\hat{i}_{Lo}(t)$, and $\hat{v}_o(t)$, $\hat{u}(t)$ is a matrix containing the control inputs $\hat{d}_1(t)$, $\hat{d}_2(t)$, and $\hat{d}_3(t)$, $\hat{y}(t)$ is a matrix containing the system outputs, and I is the identity matrix. With matrices A and B , transfer functions for PV, wind and output voltages to different duty-cycle values can be extracted according to (10). For example, $G(s)(5,1)$ represents the fifth state variable v_o and the first control variable d_1 , thus equals to open-loop transfer function of $v_o(s)/d_1(s)$. Therefore, the row number denotes the sequence of state variable, and the column number denotes that of control input

$$G = (sI - A)^{-1} B$$

$$\begin{aligned} g_{11} &= G(s)(5, 1), & g_{21} &= G(s)(1, 1), & g_{31} &= G(s)(2, 1) \\ g_{12} &= G(s)(5, 2), & g_{22} &= G(s)(1, 2), & g_{32} &= G(s)(2, 2) \\ g_{13} &= G(s)(5, 3), & g_{23} &= G(s)(1, 3), & g_{33} &= G(s)(2, 3). \end{aligned} \tag{10}$$

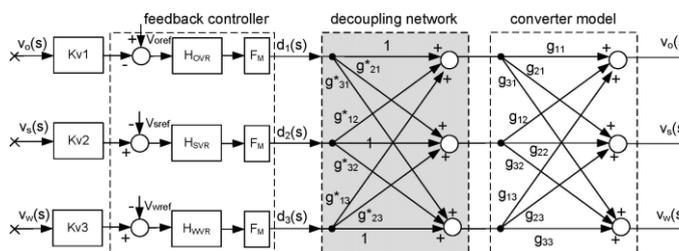


Fig. 6. Small signal model diagram, control inputs and outputs are decoupled to enable separate controller design. The far right signals are routed to the far left ones in this diagram. V_{sref} , $V_w ref$, and $V_o ref$ are the references for solar, wind and output voltages, respectively. $HSVR$, $HWVR$, and $HOVR$ are the compensators need to be designed

Fig. 6 illustrates the small signal model diagram when closing SVR, WVR, and OVR loops, which consists of the converter model and the feedback controllers. FM represents the PWM modulator gain and different K_v values represent different voltage signal sensing gains, which can be treated as the fixed proportional values.

3.3 Decoupling Method

As can be seen from Fig. 6, the three control loops are coupled with each other, which make it difficult to design close-loop compensators for each control loop.

Therefore, a decoupling network, as shadowed in Fig. 6, is introduced, so that the control loops can be designed independently with different control-loop bandwidth requirement. Since output-port voltage regulation requirement is the most stringent of the three and the PV panel and wind turbine characteristics are relatively slower, the SVR loop is designed to have a one-decade lower bandwidth than that of OVR. Moreover, WVR bandwidth can be set to be lower than that of SVR to further reduce SVR and WVR loop interactions, since the mechanical behavior of wind blades is slower than the PV behavior of PV panels. The derivation of decoupling network G^* is described as follows. The state vector matrix X can be written as $X = G \cdot U^*$, where U^* is the modified input vector made up of duty cycles U , $U^* = G^* \cdot U$. Therefore, $X = G \cdot G^* \cdot U$. According to modern control theory, our goal is to make $G \cdot G^*$ a diagonal matrix to allow one control input to determine one output independently. Therefore, based on $G^* = G^{-1} \cdot X \cdot U^{-1}$, the decoupling matrix G^* can be derived and simplified as follows:

$$G^* = \begin{bmatrix} g_{11}^* & g_{12}^* & g_{13}^* \\ g_{21}^* & g_{22}^* & g_{23}^* \\ g_{31}^* & g_{32}^* & g_{33}^* \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{g_{13}g_{23} - g_{12}g_{33}}{g_{11}g_{33} - g_{13}g_{31}} & \frac{g_{13}g_{23} - g_{13}g_{22}}{g_{11}g_{22} - g_{12}g_{21}} \\ \frac{g_{23}g_{31} - g_{21}g_{33}}{g_{11}g_{33} - g_{13}g_{31}} & 1 & \frac{g_{13}g_{21} - g_{11}g_{23}}{g_{11}g_{22} - g_{12}g_{21}} \\ \frac{g_{21}g_{32} - g_{22}g_{31}}{g_{11}g_{33} - g_{13}g_{31}} & \frac{g_{13}g_{23} - g_{12}g_{33}}{g_{11}g_{22} - g_{12}g_{21}} & 1 \end{bmatrix}$$

It should be noted that the decoupling network is only intended to calculate and derive the separate control objects, while it does not need to be implemented in the real controller design. In other word, the decoupling can be taken as one part of the control objects, but not included in the compensators. Now, the cross-coupled three-loop control system is decoupled into three independent single-loop subsystems. The system can then be controlled using independent loop controllers and each compensator can be designed separately as well. For example, the OVR

controller can then be designed based on the following plant transfer function:

$$\frac{v_o(s)}{d_1(s)} = g_{11} + g_{12} \frac{g_{23}g_{31} - g_{21}g_{33}}{g_{23}g_{33} - g_{23}g_{32}} + g_{13} \frac{g_{21}g_{32} - g_{22}g_{31}}{g_{22}g_{33} - g_{23}g_{32}} \quad (12)$$

The open-loop OVR-loop bode plot implies that it has two main poles at around LoCo resonance, which causes a -40 dB/decade slope for gain plot while not having enough phase margin. This double pole characteristic is because that this topology is buck-type derived in terms of the output port. Therefore, the design objective is to make the gain plot pass 0 dB line at -20 dB/decade slope while maintaining a sufficient phase margin. A tradition PID controller is recommended to boost the phase. The PID compensator of HOVR takes the following form:

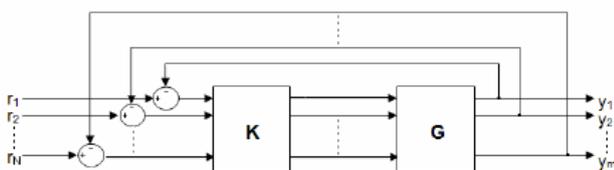
$$H_{OVR} = \frac{80(s/2\pi 400 + 1)(s/2\pi 500 + 1)}{s(s/2\pi 4000 + 1)(s/2\pi 5000 + 1)} \quad (13)$$

Similarly, SVR and WVR controllers can also be designed once their decoupled plant transfer functions are derived. The SVR and WVR bode plots before compensation have very high bandwidth. But the control bandwidth should be reduced to minimize loop interaction, SVR compensator H_{SVR} is then designed to enforce relatively low control-loop bandwidth with some phase boost. There fore, a PID controller with very low gain is adopted to achieve this design goal. And WVR compensator H_{WVR} is set at even lower gain to have a lower bandwidth than SVR loop. H_{SVR} and H_{WVR} are designed as follows: $H_{SVR} = \frac{0.08(s/2\pi 20 + 1)(s/2\pi 20 + 1)}{s(s/2\pi 1000 + 1)(s/2\pi 2000 + 1)}$ (14)

$$H_{WVR} = \frac{0.02(s/2\pi 20 + 1)(s/2\pi 30 + 1)}{s(s/2\pi 1000 + 1)(s/2\pi 1500 + 1)} \quad (15)$$

IV. INDIVIDUAL CHANNEL DESIGNING

In a typical control design task the performance is specified in terms of an output response to a given input. Meanwhile, in multivariable control, there are a number of inputs and outputs and, although it may be clear which inputs are intended to drive which outputs, the design task can be obscured by cross-coupling via the plant dynamics. Nevertheless, for clarity of both performance specification and design, it remains desirable to consider the inputs and outputs in pairs. The situation is depicted in Fig. 1, where **G** is the plant and **K** is the controller. Input *r_i* is paired with output *y_i* in accordance with specifications. An individual pairing is called a channel. Then, channel *C_i* is the pairing between *r_i* and *y_i*.



MIMO multivariable system. Channel definition

When the plant cross-coupling is weak, the design task reduces to a set of SISO design tasks and a scalar controller can be designed separately for each channel. In such context, the most appropriate methodology is to apply

classical Nyquist/Bode analysis and design to each channel ICD is a framework in which Bode/Nyquist techniques can be applied directly to the channels not only when cross-coupling is weak but in all circumstances including when cross-coupling is strong. The multivariable system is decomposed into an equivalent set of SISO systems. Each SISO system is the open-loop channel transmittance between input *r_i* and output *y_i*, with the feedback loop between output *y_i* and input *r_i* open but all other feedback loops closed, for a particular choice of *i*. What is particular to Individual Channel Design is that the SISO channel transmittances are reformulated to make explicit the role of the plant structure. Scalar multivariable structure functions (MSFs) to which the individual channel transmittances are simply related encapsulate the significant aspects of the plant structure. The multivariable nature of the original plant is maintained in the equivalent SISO systems through the multivariable structure functions with no loss of information. The ICD set up for a 2-input 2-output plant is shown next for completeness. Let a 2x2 plant

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{u}(s) \quad (1)$$

be represented by

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \quad (2)$$

where *g_{ij}(s)* represents scalar transfer functions, *y_i(s)* the outputs, and *u_i(s)* the inputs of the system, with *i, j = 1, 2*. If a

diagonal controller is given by

with *e_i(s) = r_i(s) - y_i(s)*, where *r_i(s)* represents the plant

references, then the open loop input-output channels are

clearly defined from Figs. 2 and 3 as

$$C_i(s) = k_{ii}(s)g_{ii}(s)(1 - \gamma_a(s)h_j(s)) \quad (5)$$

where *i* not equal *j* and *i, j = 1, 2*. The complex value *d* function

$$\gamma_a(s) = \frac{g_{12}(s)g_{21}(s)}{g_{11}(s)g_{22}(s)} \quad (6)$$

is referred to as the multivariable structure function (MSF).

The functions *h_i(s)* are:

$$h_i(s) = \frac{k_{ii}(s)g_{ii}(s)}{1 + k_{ii}(s)g_{ii}(s)} \quad (7)$$

The interaction or cross coupling between the channels can be evaluated through a transfer function. For instance, the influence of channel-*j* on channel-*i* is

$$d_i(s) = \frac{g_{ij}(s)}{g_{jj}(s)}h_j(s)r_j(s) \quad (8)$$

It is clear that the correct interpretation of the MSF (6) is

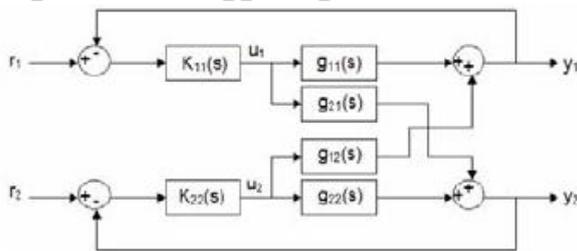
of great importance because

- It determines the dynamical characteristics of each Input–output configuration;
- It has an interpretation in the frequency domain;
- Its magnitude quantifies the coupling between the channels (in the frequency domain);
- It is related to the plant transmission zeros (zeros of $1-a(s)$), $|G(s)| = g_{11}(s)g_{22}(s) - g_{12}(s)g_{21}(s) = 0$; $a(s) = 1$ determines the non–minimum phase condition;
- Its closeness to (1,0) in the Nyquist plot indicates to
- What extent the plant is sensitive to uncertainty in terms of gain and phase margins. This fact plays a key role in order to obtain robust controllers.

A block diagram of the feedback system with the diagonal controller is shown in Fig. 2 and the equivalent scalar channels are shown in Fig. 3.

$$u(s) = K(s)e(s) \tag{3}$$

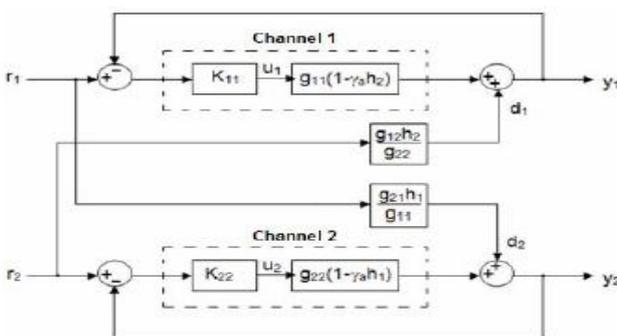
$$\begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} = \begin{bmatrix} k_{11}(s) & 0 \\ 0 & k_{22}(s) \end{bmatrix} \begin{bmatrix} e_1(s) \\ e_2(s) \end{bmatrix} \tag{4}$$



The 2-input 2-output multivariable system with a diagonal

Controller It should be emphasised that in the individual channel representation of the multivariable system there is no loss of information [4]. The multivariable character and cross coupling of the plant are contained in the MSF and the cross coupling terms. That is, (5)–(8) are equivalent to the closed loop matrix function

$$G_d(s) = (I + G(s)K(s))^{-1} G(s)K(s). \tag{9}$$



Equivalent channels of a 2-input 2-output control system It can be proven that in order to stabilise (9) it is just

necessary to stabilise the channels given by (5) [8,10]. In general stabilisation of the diagonal elements of $G(s)$ is not required [10]. The open loop system dynamical structure with a diagonal controller is summarised in Table I [4]. Notice that the coupling can be expressed in decibels directly from the channels (5) by means of functions $a(s)h_j(s)$. On the other hand, it is possible to determine the dynamical structure of the system using Table I and analysing the Nyquist plot of $(1-a(s)h_j(s))$.

Table I. Dynamical structure of open loop channels

Channel	Zeros	Poles
$C_1(s)$	Zeros of $(1-\gamma(s)h_2(s))$	Poles of $g_{11}(s), g_{12}(s), g_{21}(s), h_2(s)$
$C_2(s)$	Zeros of $(1-\gamma(s)h_1(s))$	Poles of $g_{22}(s), g_{12}(s), g_{21}(s), h_1(s)$

It is clear that the controller performance characteristics are determined by the MSF. If the transfer matrix $G(s)$ possess a non–minimum phase transmission zero, some problems will arise while stabilising it especially if the value of the zero is smaller than the desired cut–off frequency. Moreover, the robustness of the channels can be established in terms of gain and $Z = N + P$.

phase margins as the Nyquist paths of the functions $\gamma a(s)h_i(s)$ do not pass near (1,0). Thus the design of $k_{ii}(s)$, which should provide adequate gain and phase margins for $k_{ii}(s)g_{ii}(s)$, can be obtained through an iterative process. It should be noticed that the RHPPs of the channels are RHPPs of individual transfer functions as established in Table I. On the other hand, the RHPZs of the channels are RHPZs of $(1-\gamma a(s)h_i(s))$. Moreover, the number of RHPZs of the previous function can be determined after applying the Nyquist Stability Criterion. In fact, the RHPZs of $(1-\gamma a(s)h_i(s))$ are given by where P is the number of RHPPs of $a(s)h_i(s)$ and N is the number of encirclements in clockwise direction to (1,0) of the complex plane in the Nyquist diagram of $\gamma a(s)h_i(s)$. The dynamical structure of the 2x2 plant is determined by the input–output channels defined by pairing each input to each output. For instance:

- (a) $C_1(s): u_1(s)-y_1(s)$ with $\gamma_a(s) = g_{12}(s)g_{21}(s)/g_{11}(s)g_{22}(s)$
 $C_2(s): u_2(s)-y_2(s)$
- (b) $C_1(s): u_1(s)-y_2(s)$ with $\gamma_b(s) = g_{11}(s)g_{22}(s)/g_{12}(s)g_{21}(s)$
 $C_2(s): u_2(s)-y_1(s)$

The coupling characteristic of each configuration is Determined from $\gamma a(s)$ and $\gamma b(s)$ –their associated MSFs.

V. SIMULATION RESULTS

In this simulation circuit three input port are taken wind ,solar and battery sources .in this four switches are used for three switches the are give duty ratios of d1,d2 and d3.the other switch is operated with duty ratio d4 which is depend on the three switches duty ratios.

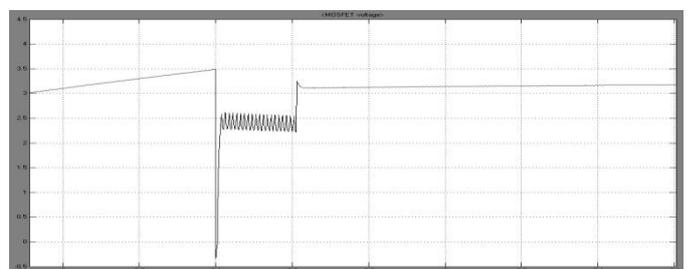


Fig.9 MOSFET voltage of S1

A four-port dc/dc converter prototype is built to verify the circuit operation. The circuit parameters are: solar port, 30– 40 V/1.5 A; wind port, 20–30 V/1.5 A; battery port, 12– 18 V/3 A; and output port, 12 V/3.3 A. The switching frequency is 100 kHz, and it is implemented by the digital control to achieve the close-loop regulation. In addition, there is no CCM and DCM transition for the output inductor current i_{Lo} , which avoids the sharp change of plant dynamic characteristics and simplifies the output-voltage feedback-controller design. The transformer magnetizing current i_p is determined by both the reflected output current and the battery current.

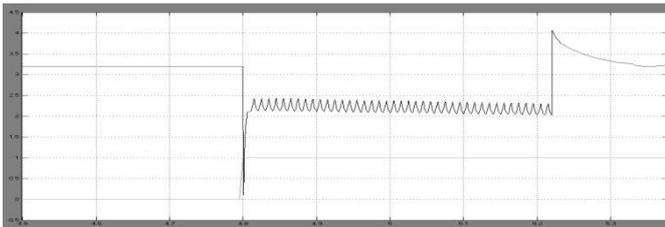


Fig.10 MOSFET voltage S1 with pulse of the switch

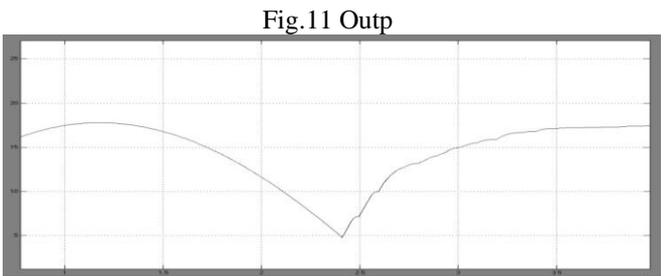


Fig.11 Output voltage

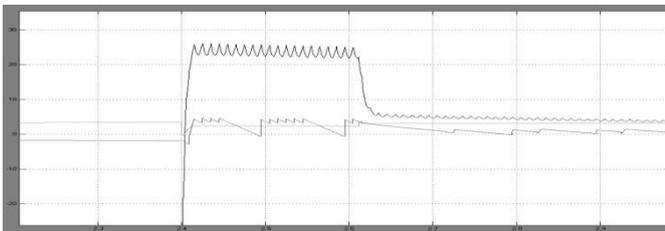


Fig.12 Output currents, pulse and voltage of switch S1

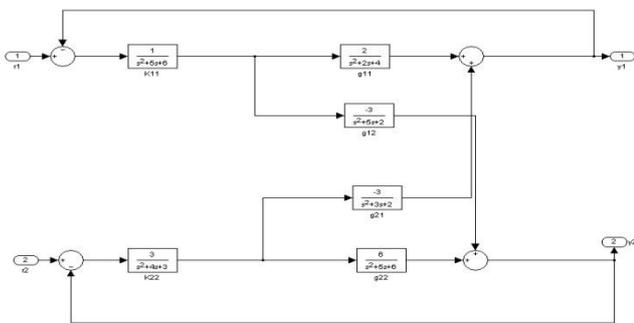


Fig.13 Icd circuit

ICD allows to coupled reference frame induction motor model. Decoupling is accomplished due to the nature of design which involves the definition individual input output

channels, which are determined multi variable function s. Classical control technique (bode and nyquist) can be used to achieve multi variable control design robust variation parameter and without ripple response. Here given below bode plots of the ICD controller in bode plots of K11, K22 are represented those will be reprocess the performance of the controller transfer functions

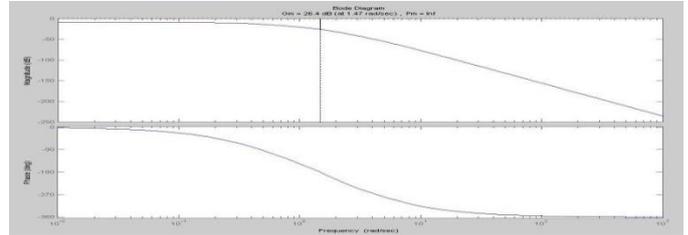


Fig.14 Bode gamma*h1

The corresponding Nyquist and Bode plots given by the toolbox are shown in Fig. 6. The user has to analyse the characteristics given in the plots, which are given next. The value of $ya(0) > 1$ (in fact, $ya(0)=2$) and the Nyquist diagram of $y a(s)$ starts at the right side of $(1,0)$. Also, $y a(s)$ has no RHPPs and its Nyquist path encircles clockwise the point $(1,0)$ twice. So, according to the Nyquist Stability Criterion, $(1- ya(s))$ contains two RHPZs. If $h2(0) > 0.5$ and stable then $(1- ya(0)h2(0)) < 0$. Thus, a stabilising controller $k11(s)$ is such that $k11(0) < 0$. On the other hand, $g11(0) > 0$; therefore, $h1(s)$ is unstable with one RHPP. Also, if $|k11(0)g11(0)| > 0$, then $h1(0) > 0$, if the relative degree of $h1(s)$ is greater than the relative degree of $ya(s)$, then $y a(s)h1(s)$ encircles counter clockwise $(1,0)$ once. Recall that $ya(s)h1(s)$ has one RHPP; thus, Channel 2 is minimum-phase. (a.4) From (a.3) it is clear that the stabilising controller of Channel 2 stabilises both $g22(s)(1-ya(s)h1(s))$ and $g22(s)$ simultaneously

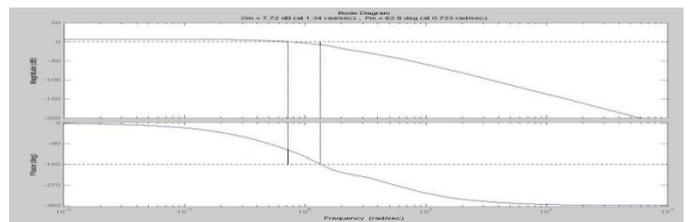


Fig.15 Bode gamma*h2

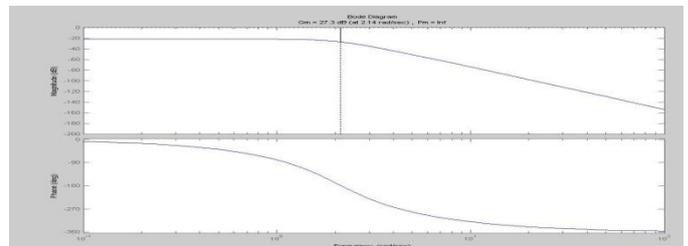


Fig.16 Bode k11

It can be concluded that the control for this configuration exists, but it presents performance limitations. the controller is designed. After performing the required iterations, the following controllers. It should be mentioned that the controller (17) satisfies the analysis carried out before. Another possible controller may exist, but the one

given before has an adequate performance (given the limitations pointed out before).

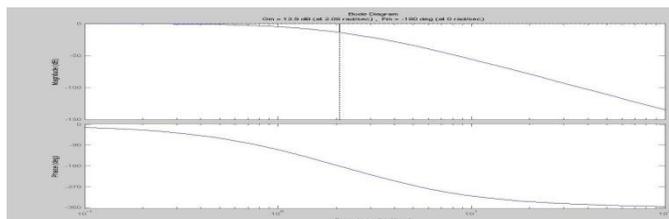


Fig.17 Bodek22

VI. CONCLUSION

In this a four port interfacing circuit is proposed which consists of the two input source ports, a bidirectional storage port, and a galvanically isolated loading port. For the four port application to the switch ZVS technique is applied and another control technique individual channel design. Modification based on the traditional half-bridge topology makes it convenient for the practicing engineers to follow the power stage design. Three degrees of freedom necessary to control power flow in the system are provided by a four-stage constant-frequency switching sequence. This four-port converter is suitable for renewable energy systems, where the energy storage is required while allowing tight load regulation.

In this paper the novel 2x2 Individual Channel Design MATLAB® Toolbox is presented. The software is a friendly programme and a valuable aid for analysing and designing multivariable control systems. The control design for a challenging plant (non minimum phase and with strong coupling) found in literature is here developed. Some designs for 2x2 MIMO systems have been actually carried out using this toolbox, proving its adequate performance.

The toolbox saves time while designing, providing all the necessary information for every design step and allowing the assessment of controller performance throughout simulations. Structural robustness and stability margins can be considered while designing, so successful controllers can be obtained considering a possible parameter variation. Some tips are given during the design process and the instructions the designer should follow are also mentioned.

This paper has presented a novel dc/dc converter topology capable of interfacing four dc power ports: two input source ports, a bidirectional storage port, and a galvanically isolated loading port. The converter features low component count and ZVS operation for all primary switches. Modification based on the traditional half-bridge topology makes it convenient for the practicing engineers to follow the power stage design. Three degrees of freedom necessary to control power flow in the system are provided by a four-stage constant-frequency switching sequence. This four-port converter is suitable for renewable energy systems, where the energy storage is required while allowing tight load regulation.

It is suitable for low-power applications since based on the half-bridge topology, while the multiport converter based on the full-bridge topology maybe suitable for high-power applications. For the hybrid PV wind system, the proposed control structure is able to achieve maximum power harvesting for PV and/or wind power sources,

meanwhile maintaining a regulated output voltage. The close-loop controller design is investigated based on the dynamic modelling of the converter power stage. Proper decoupling method is introduced to help design close-loop compensators for such a cross-coupled control system. The circuit operation of this converter and its control system is experimentally verified. Although the proposed four port converter only has two input ports, it can be extended to have n input ports. In this a control technique called ICD used for getting good optimization results of system; Individual channel designing is one of the best control techniques in upcoming control techniques. This individual channel designing implemented for 2X2 matrix .because of this one we can get out which is best output.

REFERENCES

- [1] Z. Qian, O. Abdel-Rahman, and I. Batarseh, "An Integrated Four-Port DC/DC Converter for Renewable Energy Applications, IEEE Trans. Power Electron., vol. 25, no. 7, pp. 1877–1887, JULY. 2010.
- [2] Y. Liu and Y. M. Chen, "A systematic approach to synthesizing multi input DC–DC converters," IEEE Trans. Power Electron., vol. 24, no. 2, pp. 116–127, Jan. 2009.
- [3] B. G. Dobbs and P. L. Chapman, "A multiple-input DC-DC converter topology," in Proc. IEEE Power Electron. Lett., Mar, 2003, vol. 1, pp. 6–9.
- [4] N. D. Benavides and P. L. Chapman, "Power budgeting of a multiple input buck-boost converter," IEEE Trans. Power Electron., vol. 20, no. 6, pp. 1303–1309, Nov. 2005.
- [5] H. Matsuo, W. Lin, F. Kurokawa, T. Shigemizu, and N. Watanabe, "Characteristics of the multiple-input DC–DC converter," IEEE Trans. Ind. Appl., vol. 51, no. 3, pp. 625–631, Jun. 2004.
- [6] A. Khaligh, J. Cao, and Y. Lee, "A multiple-input DC–DC converter
- [7] Y. M. Chen, Y. C. Liu, and F. Y. Wu, "Multi-input DC/DC converter based on the multiwinding transformer for renewable energy applications," IEEE Trans. Ind. Appl., vol. 38, no. 4, pp. 1096–1104, Aug. 2002.
- [8] A. Kwasinski, "Identification of feasible topologies for multiple-input DCDC converters," IEEE Trans. Power Electron., vol. 24, no. 3, pp. 856–861, Mar. 2009.
- [9] G. Su and L. Tang, "A reduced-part, triple-voltage DC-DC converter for EV/HEV power management," IEEE Trans. Power Electron., vol. 24, no. 10, pp. 2406–2410, Oct. 2009.
- [10] F. Z. Peng, H. Li, G. J. Su, and J. S. Lawler, "A new ZVS bidirectional DC-DC converter for fuel cell and battery applications," IEEE Trans. Power Electron., vol. 19, no. 1, pp. 54–65, Jan. 2004.
- [11] H. Tao, J. L. Duarte, and M. A. M. Hendrix, "Three-port triple-half-bridge bidirectional converter with zero-voltage switching," IEEE Trans. Power Electron., vol. 23, no. 2, pp. 782–792, Mar. 2008.
- [12] W. Jiang and B. Fahimi, "Multi-port power electric interface for renewable energy sources," in Proc. IEEE Appl. Power Electron. Conf., 2009, pp. 347–352.
- [13] D. Liu and H. Li, "A ZVS bi-directional DC-DC converter for multiple energy storage elements," IEEE

- Trans. Power Electron., vol. 21, no. 5, pp. 1513–1517, Sep. 2006.
- [14] C. Zhao, S.D.Round, and J.W.Kolar, “An isolated three-port bidirectional DC-DC converter with decoupled power flow management,” IEEE Trans. Power Electron., vol. 21, no. 5, pp. 2443–2453, Sep. 2008.
- [15] H. Tao, A.Kotsopoulos, J. L.Duarte, andM.A.M.Hendrix, “Transformer-coupled multiport ZVS bidirectional DC–DC converter with wide input range,” IEEE Trans. Power Electron., vol. 23, pp. 771–781, Mar. 2008.
- [16] J. L. Duarte, M. Hendrix, and M. G. Simoes, “Three-port bidirectional converter for hybrid fuel cell systems,” IEEE Trans. Power Electron., vol. 22, no. 2, pp. 480–487, Mar. 2007.
- [17] H. Al-Atrash and I. Batarseh, “Boost-integrated phase-shift full-bridge converters for three-port interface,” in Proc. IEEE Power Electron. Spec. Conf., 2007, pp. 2313–2321.
- [18] H. Tao, A. Kotsopoulos, J. L. Duarte, and M. A. M. Hendrix, “Family of multiport bidirectional DC-DC converters,” in Proc. IEEE Power Electron. Spec. Conf., 2008, pp. 796–801.
- [19] H. Al-Atrash, F. Tian, and I. Batarseh, “Tri-modal half-bridge converter topology for three-port interface,” IEEE Trans. Power Electron., vol. 22, no. 1, pp. 341–345, Jan. 2007.
- [20] Z. Qian, O. Abdel-Rahman, J. Reese, H. Al-Atrash, and I. Batarseh, “Dynamic analysis of three-port DC/DC converter for space applications,” in Proc. IEEE Appl. Power Electron. Conf., 2009, pp. 28–34.
- [21] Z. Qian, O. Abdel-Rahman, M. Pepper, and I. Batarseh, “Analysis and design for paralleled three-port DC/DC converters with democratic current sharing control,” in Proc. IEEE Energy Convers. Congr. Expo., 2009, pp. 1375–1382.
- [22] H. Mao, J. Abu-Qahouq, S. Luo, and I. Batarseh, “Zero-voltage-switching half-bridge DC-DC converter with modified PWM control method,” IEEE Trans. Power Electron., vol. 19, no. 4, pp. 947–958, Jul. 2004.
- [23] H. Tao, A. Kotsopoulos, J. L. Duarte, and M. A. M. Hendrix, “Multiinput bidirectional DC-DC converter combining DC-link and magnetic coupling for fuel cell systems,” in Proc. IEEE Ind. Appl. Conf., 2005, pp. 2021–2028.
- [24] R. W. Erickson and D. Maksimovic, Fundamentals of Power Electronics, 2nd ed. Boston, MA: Kluwer, 2000, ch. 7.
- [25] H. Al-Atrash, M. Pepper, and I. Batarseh, “A zero-voltage switching three-port isolated full-bridge converter,” in Proc. IEEE Int. Telecommun. Energy Conf., 2006, pp.411–418.
- [26] Z. Qian, O. Abdel-Rahman, H. Hu, and I. Batarseh, “Multi-channel threeport DC/DC converters as maximal power tracker, battery charger and bus regulator,” in Proc. IEEE Appl. Power Electron. Conf., 2010, pp. 2073–2079.
- [27] H. Krishnaswami and N. Mohan, “Three-port series-resonant DC–DC converter to interface renewable energy sources with bidirectional load and energy storage ports,” IEEE Trans. Power Electron., vol. 24, no. 9–10, pp. 2289–2297, Oct. 2009.
- [28] Z. Qian, O. Abdel-Rahman, H. Al-trash, and I. Batarseh, “Modeling and control of three-port DC/DC converter interface for satellite applications, IEEE Trans. Power Electron., vol. 25, no. 3, pp. 637–649, Mar. 2010.

A Ticket Based Architecture for Providing Anonymity and Traceability in Wireless Mesh Network

S.Parvaiz¹, V. Sreenatha Sarma²

^{1,2}(M. tech, Department of CSE, ASCET, GUDUR,

ABSTRACT: A wireless mesh is a communication network made up of radio nodes organized in a mesh topology. A wireless mesh network help users to stay online anywhere, anytime for an unlimited time and it provide high security. In this paper we come across two issues one is anonymity and other is traceability. Anonymity provides protection for users to enjoy network services without being traced. Now a day's Anonymity has received increasing attention in the literature due to the users' awareness of their privacy. Anonymity-related issues have been extensively studied in payment-based systems such as e-cash and peer-to-peer (P2P) systems, little effort has been devoted to wireless mesh networks (WMNs). The network authority requires conditional anonymity such that misbehaving entities in the network remain traceable. so in order to provide high network security we provide a secure architecture to ensure unconditional anonymity for honest users and traceability of misbehaving users for network authorities in WMNs. This architecture strives to resolve the conflicts between the anonymity and traceability objectives. Finally the architecture guaranteeing fundamental security requirements such as authentication, confidentiality, data integrity, and nonrepudiation.

I. INTRODUCTION

Traceability is the ability to map events in cyberspace, particularly on the Internet, back to real-world instigators, often with a view to holding them accountable for their actions. Anonymity is present when traceability fails. Failures of traceability, with consequent unintentional anonymity, have continued as the technology has changed. The underlying reason for this continuing failure is a lack of economic incentives for improvement. The lack of traceability at the edges is further illustrated by a new method of stealing another person's identity on an Ethernet Local Area Network that existing tools and procedures would entirely fail to detect. Anonymity and privacy issues have gained considerable research efforts, which have focused on investigating anonymity in different context or application scenarios. One requirement for anonymity is to unlink a user's identity to his or her specific activities, such as the anonymity fulfilled in the untraceable e-cash systems and the P2P Payment systems, where the payments cannot be linked to the identity of a payer by the bank or broker.

Anonymity is also required to hide the location information of a user to prevent movement tracing, as is important in mobile networks and VANETs. In wireless communication systems, it is easier for a global observer to mount traffic analysis attacks by following the packet forwarding path than in wired networks. Thus, routing anonymity is indispensable, which conceals the confidential communication relationship of two parties by building an anonymous path between them. Nevertheless, unconditional anonymity may incur insider attacks since misbehaving

users are no longer traceable. Therefore, traceability is highly desirable such as in e-cash systems, where it is used for detecting and tracing double-spenders. A wireless mesh network (WMN) is a communications network made up of radio nodes organized in a mesh topology. Wireless mesh networks often consist of mesh clients, mesh routers and gateways. The mesh clients are often laptops, cell phones and other wireless devices while the mesh routers forward traffic to and from the gateways. A mesh network is reliable and offers redundancy. When one node can no longer operate, the rest of the nodes can still communicate with each other, directly or through one or more intermediate nodes. A Wireless mesh networks can be implemented with various wireless technology including 802.11, 802.15, 802.16, cellular technologies or combinations of more than one type. Wireless mesh network can be seen as a special type of wireless ad-hoc network. Wireless Mesh Networks (WMNs) have become the focus of much research since they allow for increased coverage while retaining the attractive features of low cost and easy deployment. WMNs have been identified as key technology to enhance and compliment existing network installations as well as provide access where traditional technology is not available or too costly to install. A WMN is made up of mesh routers (MRs), which have limited or no mobility, and mesh clients (MCs) which are often fully mobile. The mesh routers form the backbone of the network allowing the clients to have access to the network through the backbone. We propose an algorithm for fair scheduling in WMNs with multiple gateways. We also propose another algorithm for scheduling which places more emphasis on throughput while retaining a basic level of throughput called mixed-bias. This technique biases against characteristics of the network which are detrimental to performance, fairness, or both. Many protocols currently implemented for WMNs have evolved from traditional single-hop wireless local area networks (WLAN) and mobile ad-hoc networks (MANET). However, both of these networks have characteristics which make them very different from WMNs. While WLANs have relatively static topologies, MANETs on the other hand are fully mobile.

Therefore, using protocols designed solely for either of these networks alone does not take advantage of some of the most advantageous features of WMNs. In MANETs all nodes are routers and suffer from limited power and bandwidth. In a WMN the MRs have greater resources available than the MCs which is a property that may be exploited. Although a lot of research efforts have been made to address these problems and some new specialized algorithms have been proposed specifically for WMNs, there are still many challenges in the area. Many of the existing solutions make many assumptions that can be relaxed to allow for a more general approach to be taken.

Here, we are motivated by resolving the above security conflicts, namely anonymity and traceability, in the

emerging WMN communication systems. We have proposed the initial design of our security architecture, where the feasibility and applicability of the architecture were not fully understood. As a result, we provide detailed efficiency analysis in terms of storage, communication, and computation in this paper to show that our SAT is a practically viable solution to the application scenario of interest. Our system borrows the blind signature technique from payment systems and hence, can achieve the anonymity of unlinking user identities from activities, as well as the traceability of misbehaving users. Furthermore, the proposed pseudonym technique renders user location information unexposed. Our work differs from previous work in that WMNs have unique hierarchical topologies and rely heavily on wireless links, which have to be considered in the anonymity design. As a result, the original anonymity scheme for payment systems among bank, customer, and store cannot be directly applied. In addition to the anonymity scheme, other security issues such as authentication, key establishment, and revocation are critical in WMNs to ensure the correct application of the anonymity scheme. Moreover, although we employ the widely used pseudonym approach to ensure network access anonymity and location privacy, our pseudonym generation does not rely on a central authority.

II. System Architecture

A large number of studies on multi-hop wireless networks have been devoted to system stability while maximizing metrics like throughput or utility. These metrics measure the performance of a system over a long time-scale. For a large class of applications such as video or voice over IP, embedded network control and for system design; metrics like delay are of prime importance. The delay performance of wireless networks, however, has largely been an open problem. This problem is notoriously difficult even in the context of wire line networks, primarily because of the complex interactions in the network (e.g., superposition, routing, departure, etc.) that make its analysis amenable only in very special cases like the product form networks. The problem is further exacerbated by the mutual interference inherent in wireless networks which, complicates both the scheduling mechanisms and their analysis. Some novel analytical techniques to compute useful lower bound and delay estimates for wireless networks with single hop traffic were developed.

We analyze a multi-hop wireless network with multiple source-destination pairs, given routing and traffic information. Each source injects packets in the network, which traverses through the network until it reaches the destination. For example, a multi-hop wireless network with three flows. The exogenous arrival processes correspond to the number of packets injected in the system at time. A packet is queued at each node in its path where it waits for an opportunity to be transmitted. Since the transmission medium is shared, concurrent transmissions can interfere with each others' transmissions. The set of links that do not cause interference with each other can be scheduled simultaneously, and we call them *activation vectors* (matchings). We do not impose any a priori restriction on the set of allowed activation vectors, i.e., they can characterize any combinatorial interference model. For example, in a K-hop interference model, the links scheduled

simultaneously are separated by at least K hops. Each link has unit capacity; i.e., at most one packet can be transmitted in a slot. For the above example, we assume a 1-hop interference model. The delay performance of any scheduling policy is primarily limited by the interference, which causes many bottlenecks to be formed in the network. We demonstrated the use of exclusive sets for the purpose of deriving lower bounds on delay for a wireless network with single hop traffic.

A. Proposed approach

All projects are feasible when provided with unlimited resources and infinite time! Unfortunately, the development of computer-based system or product is more likely plagued by a scarcity of resources and difficult delivery dates. It is both necessary and prudent to evaluate the feasibility of a project at the earliest possible time. Months or years of effort, thousands or millions of dollars, and untold professional embarrassment can be averted if an ill-conceived system is recognized early in the definition phase. Feasibility and risk analysis are related in many ways. If project risk is great the feasibility of producing quality software is reduced. During product engineering, however, we concentrate our attention on four primary areas of interest.

B. Technical Feasibility

This application in going to be used in an Internet environment called www (World Wide Web). So, it is necessary to use a technology that is capable of providing the networking facility to the application. This application as also able to work on distributed environment. Application on developed with J2EE (Java 2 Enterprise Edition platform) Technology. One major advantage in application is platform neutral. We can deploy and used it in any operating system. GUI is developed using HTML to capture the information from the customer. HTML is used to display the content on the browser. It uses TCP/IP protocol. It is an interpreted language. It is very easy to develop a page/document using HTML some RAD (Rapid Application Development) tools are provided to quickly design/develop our application. So many objects such as button, text fields, and text area etc are provided to capture the information from the customer.

C. Economical Feasibility

The economical issues usually arise during the economical feasibility stage are whether the system will be used if it is developed and implemented, whether the financial benefits are equal and exceeds the costs. The cost for developing the project will include cost conducts full system investigation, cost of hardware and software for the class of being considered, the benefits in the form of reduced costs or fewer costly errors. The project is economically feasible if it is developed and installed. It reduces the work load. Keep the class of application in the view, the cost of hardware and software is considered to be economically feasible.

D. Operational Feasibility

In our application front end is developed using GUI. So it is very easy to the customer to enter the necessary information. But customer has some knowledge

on using web applications before going to use our application.

E. Social Feasibility

It is a determination of whether the people will accept a proposed project or not.

F. Management Feasibility

It determines whether the proposed project will be acceptable to the management.

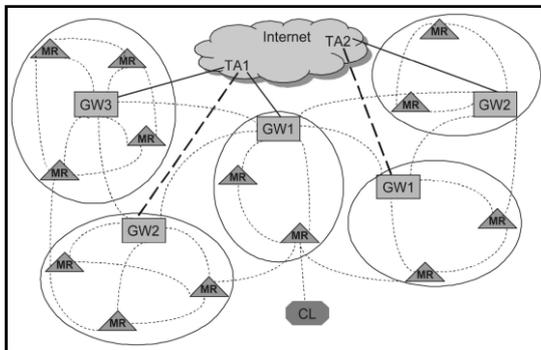
G. Legal Feasibility

These concerns about the legalities are satisfied.

H. Time Feasibility

It determines whether a proposed project can be implemented fully within stipulated time. We strongly feel that the proposed system is feasible in all respects.

Network Architecture:



The wireless mesh backbone consists of mesh routers (MRs) and gateways (GWs) interconnected by ordinary wireless links (shown as dotted curves). Mesh routers and gateways serve as the access points of the WMN and the last resorts to the Internet, respectively. Each WMN domain, or trust domain (to be used interchangeably) is managed by a domain administrator that serves as a trusted authority (TA), e.g., the central server of a campus WMN. TAs and gateways are assumed to be capable of handling computationally intensive tasks. In addition, they are assumed to be protected in private places and cannot be easily compromised due to their important roles in the WMN.

III. TECHNIQUES OF ARCHITECTURE

A. Wireless mesh networks (WMNs)

The wireless mesh backbone consists of mesh routers (MRs) and gateways (GWs) interconnected by ordinary wireless links (shown as dotted curves). Mesh routers and gateways serve as the access points of the WMN and the last resorts to the Internet, respectively. Each WMN domain, or trust domain (to be used interchangeably) is managed by a domain administrator that serves as a trusted authority the central server of a campus WMN.

B. Blind Signature

In general, a blind signature scheme allows a receiver to obtain a signature on a message such that both the message and the resulting signature remain unknown to

the signer. We refer the readers for a formal definition of a blind signature scheme, which should bear the properties of verifiability, unlinkability, and unforgeability. Blind signature scheme, where the restrictiveness property is incorporated into the blind signature scheme such that the message being signed must contain encoded information. As the name suggests, this property restricts the user in the blind signature scheme to embed some account-related secret information into what is being signed by the bank (otherwise, the signing will be unsuccessful) such that this secret can be recovered by the bank to identify a user if and only if he double-spends. The restrictiveness property is essentially the guarantee for traceability in the restrictive blind signature systems.

C. Ticket Issuance

In order to maintain security of the network against attacks and the fairness among clients, the home server manager may control the access of each client by issuing tickets based on the misbehavior history of the client, which reflects the server manager’s confidence about the client to act properly. Ticket issuance occurs when the client initially attempts to access the network or when all previously issued tickets are depleted. The client needs to reveal his real ID to the server manager in order to obtain a ticket since the server manager has to ensure the authenticity of this client.

D. Fraud Detection

Fraud is used interchangeably with misbehavior in this paper, which is essentially an insider attack. Ticket reuse generally results from the client’s inability to obtain tickets from the TA when network access is desired, primarily due to the client’s past misbehavior, which causes the server manager to constrain his ticket requests.

E. Fundamental security objectives

It is trivial to show that our security architecture satisfies the security requirements for authentication, data integrity, and confidentiality, which follows directly from the employment of the standard cryptographic primitives, message authentication code, and encryption, in our system. We are only left with the proof of nonrepudiation in this category. A fraud can be repudiated only if the client can provide a different representation, he knows of message from what is derived by the server manager. If the client has misbehaved, the representation he knows will be the same as the one derived by the server Manager which ensures nonrepudiation. Ad hoc networks inherit some of the traditional problems of wireless communication and wireless networking:

- I. The wireless medium does not have proper boundaries outside of which nodes are known to be unable to receive network frames.
- II. The wireless channel is weak, unreliable, and unprotected from outside signals, which may cause lots of problems to the nodes in the network.
- III. The wireless channel has time-varying and asymmetric propagation properties.
- IV. Hidden-node and exposed-node problems may occur.

IV. Result

We are motivated by resolving the above security

conflicts, namely anonymity and traceability, in the emerging WMN communication systems. We have proposed the initial design of our security architecture, where the feasibility and applicability of the architecture were not fully understood. As a result, we provide detailed efficiency analysis in terms of storage, communication, and computation in this paper to show that our SAT is a practically viable solution to the application scenario of interest. Our system borrows the blind signature technique from payment systems, and hence, can achieve the anonymity of unlinking user identities from activities, as well as the traceability of misbehaving users. Furthermore, the proposed pseudonym technique renders user location information unexposed. Our work differs from previous work in that WMNs have unique hierarchical topologies and rely heavily on wireless links, which have to be considered in the anonymity design. As a result, the original anonymity scheme for payment systems among bank, customer, and store cannot be directly applied. In addition to the anonymity scheme, other security issues such as authentication, key establishment, and revocation are critical in WMNs to ensure the correct application of the anonymity scheme. Moreover, although we employ the widely used pseudonym approach to ensure network access anonymity and location privacy, our pseudonym generation does not rely on a central authority, e.g., the broker, the domain authority, the transportation authority or the manufacturer, and the trusted authority, who can derive the user's identity from his pseudonyms and illegally trace an honest user. Our system is not intended for achieving routing anonymity, which can be incorporated as an enhancement. So, Finally the architecture guaranteeing fundamental security requirements such as authentication, confidentiality, data integrity, and nonrepudiation.

V. CONCLUSION

We propose SAT, a security architecture mainly consisting of the ticket-based protocols, which resolves the conflicting security requirements of unconditional anonymity for honest users and traceability of misbehaving users. By utilizing the tickets, self-generated pseudonyms, and the hierarchical identity-based cryptography, the proposed architecture is demonstrated to achieve desired security objectives and efficiency. In the WMNs considered here, the uplink from the client to the mesh router may rely on multihop communications. Peer clients act as relaying nodes to forward each other's traffic to the mesh router, which forms a P2P network. The notorious problem common in P2P communication systems is the free-riding, where some peers take advantage of the system by providing little or no service to other peers or by leaving the system immediately after the service needs are satisfied. Peer cooperation is thus the fundamental requirement for P2P systems to operate properly. Since peers are assumed to be selfish, incentive mechanisms become essential to promote peer cooperation in terms of both cooperativeness and availability. Typical incentive mechanisms for promoting cooperativeness include reputation and payment-based approaches. In the reputation-based systems, peers are punished or rewarded based on the observed behavior. However, low availability remains an unobservable behavior in such systems, which hinders the feasibility of the reputation-based mechanism in improving peer

availability. By contrast, the payment-based approach provides sufficient incentives for enhancing both cooperativeness and availability, and thus, is ideal to be employed in multihop uplink communications among peer clients in our WMN system.

REFERENCES

- [1] European Telecomm. Standards Inst. (ETSI), "GSM 2.09: Security Aspects," June 1993.
- [2] P. Kyasanur and N.H. Vaidya, "Selfish MAC Layer Misbehavior in Wireless Networks," IEEE Trans. Mobile Computing, vol. 4, no. 5, pp. 502-516, Sept. 2005.
- [3] A. Perrig, J. Stankovic, and D. Wagner, "Security in Wireless Sensor Networks," Comm. ACM, vol. 47, no. 6, pp. 53-57, 2004.
- [4] S. Zhu, S. Setia, and S. Jajodia, "LEAP+: Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks," ACM Trans. Sensor Networks, vol. 2, no. 4, pp. 500-528, Nov. 2006.
- [5] W. Lou and Y. Fang, A Survey on Wireless Security in Mobile Ad Hoc Networks: Challenges and Possible Solutions, X. Chen, X. Huang, and D.-Z. Du, eds., Kluwer Academic Publishers/Springer, 2004.
- [6] L. Zhou and Z.J. Haas, "Securing Ad Hoc Networks," IEEE Network Magazine, vol. 13, no. 6, pp. 24-30, Dec. 1999.
- [7] M. Raya and J-P. Hubaux, "Securing Vehicular Ad Hoc Networks," J. Computer Security, special issue on security of ad hoc and sensor networks, vol. 15, no. 1, pp. 39-68, 2007.
- [8] N.B. Salem and J-P. Hubaux, "Securing Wireless Mesh Networks," IEEE Wireless Comm., vol. 13, no. 2, pp. 50-55, Apr. 2006.
- [9] Y. Zhang and Y. Fang, "ARSA: An Attack-Resilient Security Architecture for Multihop Wireless Mesh Networks," IEEE J. Selected Areas Comm., vol. 24, no. 10, pp. 1916-1928, Oct. 2006.
- [10] I.F. Akyildiz, X. Wang, and W. Wang, "Wireless Mesh Networks: A Survey," Computer Networks, vol. 47, no. 4, pp. 445-487, Mar. 2005.



S. Parvaiz was born in proddatur, A.P, India. He received the B. Tech (IT) degree from the Madina Engineering College, kadapa in 2010: and Pursuing M.Tech (Computer Science) from the Audisankara college of engineering and technology, Gudur.

Power Flow Control by Using DPFC

T. Obulesu¹, S. Sarada², M. Sudheer babu³

^{1,3}M.Tech Student, Department of EEE A.I.T.S Engineering College Kadapa, India

²Associative professor of AITS college

Abstract: this paper presents a new component within the flexible ac-transmission system (FACTS) family, called distributed power-flow controller (DPFC). The DPFC is derived from the unified power-flow controller (UPFC). The DPFC can be considered as a UPFC with an eliminated common dc link. The active power exchange between the shunt and series converters, which is through the common dc link in the UPFC, is now through the transmission lines at the third-harmonic frequency. The DPFC employs the distributed FACTS (D-FACTS) concept, which is to use multiple small-size single-phase converters instead of the one large-size three-phase series converter in the UPFC. The large number of series converters provides redundancy, thereby increasing the system reliability. As the D-FACTS converters are single-phase and floating with respect to the ground, there is no high-voltage isolation required between the phases. Accordingly, the cost of the DPFC system is lower than the UPFC. The DPFC has the same control capability as the UPFC, which comprises the adjustment of the line impedance, the transmission angle, and the bus voltage. The principle and analysis of the DPFC are presented in this paper and the corresponding experimental results that are carried out on a scaled prototype are also shown.

Index Terms: AC–DC power conversion, load flow control, power electronics, power semiconductor devices, power system control, power-transmission control.

I. INTRODUCTION

THE GROWING demand and the aging of networks make it desirable to control the power flow in power-transmission systems fast and reliably. The flexible ac-transmission system (FACTS) that is defined by IEEE as “a power-electronic-based system and other static equipment that provide control of one or more ac-transmission system parameters to enhance controllability and increase power-transfer capability” [2], and can be utilized for power-flow control. Currently, the unified power-flow controller (UPFC) shown in Fig. 1, is the most powerful FACTS device, which can simultaneously control all the parameters of the system: the line impedance, the transmission angle, and bus voltage.

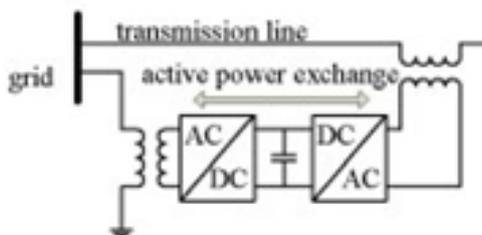


Fig. 1. Simplified representation of a UPFC.

The UPFC is the combination of a static synchronous compensator (STATCOM) and a static synchronous series compensator (SSSC), which are coupled via a common dc link, to allow bidirectional flow of active power between the series output terminals of the SSSC and the shunt output terminals of the STATCOM. The converter in series with the line provides the main function of the UPFC by injecting a four-quadrant voltage with controllable magnitude and phase. The injected voltage essentially acts as a synchronous ac-voltage source, which is used to vary the transmission angle and line impedance, thereby independently controlling the active and reactive power flow through the line. The series voltage results in active and reactive power injection or absorption between the series converter and the transmission line. This reactive power is generated internally by the series converter (see e.g., SSSC), and the active power is supplied by the shunt converter that is back-to-back connected. The shunt converter controls the voltage of the dc capacitor by absorbing or generating active power from the bus; therefore, it acts as a synchronous source in parallel with the system. Similar to the STATCOM, the shunt converter can also provide reactive compensation for the bus.

The components of the UPFC handle the voltages and currents with high rating; therefore, the total cost of the system is high. Due to the common dc-link interconnection, a failure that happens at one converter will influence the whole system. To achieve the required reliability for power systems, bypass circuits and redundant backups (backup transformer, etc.) are needed, which on other hand, increase the cost. Accordingly, the UPFC has not been commercially used, even though, it has the most advanced control capabilities.

This paper introduces a new concept, called distributed power-flow controller (DPFC) that is derived from the UPFC. The same as the UPFC, the DPFC is able to control all system parameters. The DPFC eliminates the common dc link between the shunt and series converters. The active power exchange between the shunt and the series converter is through the transmission line at the third-harmonic frequency. The series converter of the DPFC employs the distributed FACTS

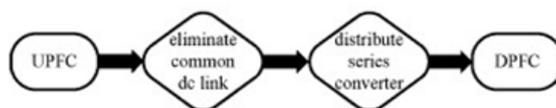
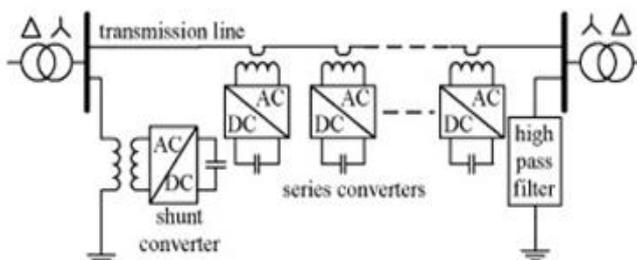


Fig. 2. Flowchart from UPFC



to DPFC.

Fig. 3. DPFC configuration.

(D-FACTS) concept . Comparing with the UPFC, the DPFC have two major advantages: 1) low cost because of the low voltage isolation and the low component rating of the series converter and 2) high reliability because of the redundancy of the series converters. This paper begins with presenting the principle of the DPFC, followed by its steady-state analysis. After a short introduction of the DPFC control, the paper ends with the experimental results of the DPFC.

II. DPFC PRINCIPLE

Two approaches are applied to the UPFC to increase the reliability and to reduce the cost; they are as follows. First, eliminating the common dc link of the UPFC and second distributing the series converter, as shown in Fig. 2. By combining these two approaches, the new FACTS device—DPFC is achieved.

The DPFC consists of one shunt and several series-connected converters. The shunt converter is similar as a STATCOM, while the series converter employs the D-FACTS concept, which is to use multiple single-phase converters instead of one large rated converter. Each converter within the DPFC is independent and has its own dc capacitor to provide the required dc voltage. The configuration of the DPFC is shown in Fig. 3.

As shown, besides the key components, namely the shunt and series converters, the DPFC also requires a high-pass filter that is shunt connected at the other side of the transmission line, and two Y-Δ transformers at each side of the line. The reason for these extra components will be explained later.

The unique control capability of the UPFC is given by the back-to-back connection between the shunt and series converters, which allows the active power to exchange freely. To ensure that the DPFC have the same control capability as the UPFC, a method that allows the exchange of active power between converters with eliminated dc link is the prerequisite.

A. Eliminate DC Link

Within the DPFC, there is a common connection between the ac terminals of the shunt and the series converters, which is the transmission line. Therefore, it is possible to exchange the active power through the ac terminals of the converters. The method is based on the power theory of nonsinusoidal components. According to the Fourier analysis, a nonsinusoidal voltage and current can be expressed by the sum of sinusoidal functions in different frequencies with different amplitudes. The active power resulting from this nonsinusoidal voltage and current is defined as the mean value of the product of voltage and current.

cross product of terms with

Since the integrals of all the different

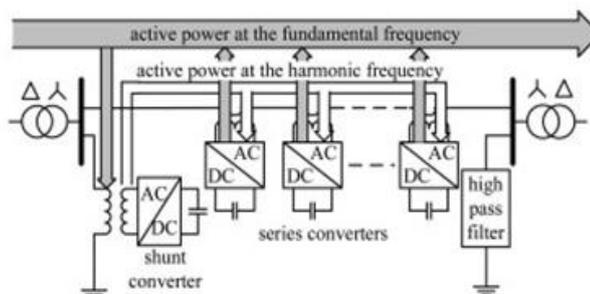


Fig. 4. Active power exchange between DPFC converters.

Frequencies are zero; the active power can be expressed by

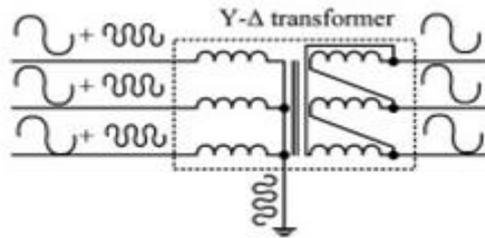
$$P = \sum_{i=1}^{\infty} V_i I_i \cos \phi_i \quad (1)$$

Where V_i and I_i are the voltage and current at the i th harmonic frequency, respectively, and ϕ_i is the corresponding angle between the voltage and current. Equation (1) describes that the active power at different frequencies is isolated from each other and the voltage or current in one frequency has no influence on the active power at other frequencies. The independency of the active power at different frequencies gives the possibility that a converter without power source can generate active power at one frequency and absorb this power from other frequencies.

By applying this method to the DPFC, the shunt converter can absorb active power from the grid at the fundamental frequency and inject the current back into the grid at a harmonic frequency. This harmonic current will flow through the transmission line. According to the amount of required active power at the fundamental frequency, the DPFC series converters generate a voltage at the harmonic frequency, thereby absorbing the active power from harmonic components. Assuming a lossless converter, the active power generated at fundamental frequency is equal to the power absorbed from the harmonic frequency. For a better understanding, Fig. 4 indicates how the active power exchanges between the shunt and the series converters in the DPFC system.

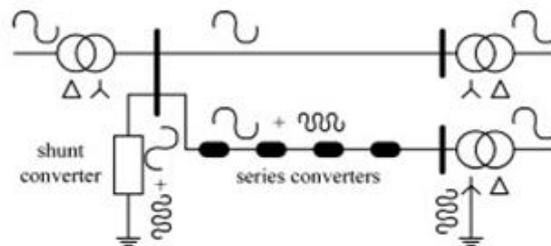
The high-pass filter within the DPFC blocks the fundamental frequency components and allows the harmonic components to pass, thereby providing a return path for the harmonic components. The shunt and series converters, the high-pass filter, and the ground form the closed loop for the harmonic current.

Due to the unique characters of third-harmonic frequency components, the third harmonic is selected to exchange the active power in the DPFC. In a three-phase system, the third harmonic in each phase is identical, which is referred to as "zero-sequence." The zero-sequence harmonic can be naturally blocked by Y-Δ transformers, which are widely used in power system to change voltage level. Therefore, there is no extra filter required to prevent the harmonic leakage to the rest of the network. In addition, by using the third harmonic, the costly high-pass filter, as shown in Fig. 4, can be replaced by a cable that is connected between transformer on the right side in Δ winding appears open circuit to harmonic current will flow through the grounding cable, as shown in high-pass filter is eliminated.



the neutral point of the Y-Δ Fig. 3 and the ground. Because the third-harmonic current, all the Y-winding and concentrate to Fig. 5. Therefore, the large-size

Fig. 5. Utilize grounded Y-Δ for the zero-sequence third harmonic.



transformer to provide the path

Fig. 6. Route the harmonic current by using the grounding status of the Y-Δ transformer.

Another advantage of using third harmonic to exchange activepower is that the way of grounding of Y-Δ transformers can be used to route the harmonic current in a meshed network. If the branch requires the harmonic current to flow through, the neutral point of the Y-Δ transformer at the other side in that branch will be grounded and vice versa. Fig. 6 demonstrates a simple example of routing the harmonic current by using a grounding Y-Δ transformer. Because the transformer of the line without the series converter is floating, it is open circuit for third-harmonic components. Therefore, no third-harmonic current will flow through this line.

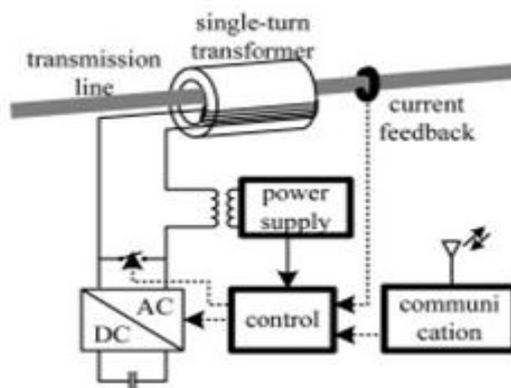
Theoretically, the third-, sixth-, and ninth-harmonic frequencies are all zero-sequence, and all can be used to exchange active power in the DPFC. As it is well known, the capacity of a transmission line to deliver power depends on its

impedance. Since the transmission-line impedance is inductive and proportional to the frequency, high-transmission frequencies will cause high impedance. Consequently, the zero-sequence harmonic with the lowest frequency—third harmonic is selected.

B. Distributed Series Converter

The D-FACTS is a solution for the series-connected FACTS, which can dramatically reduce the total cost and increase the reliability of the series FACTS device. The idea of the D-FACTS is to use a large number of controllers with low rating instead of one large rated controller. The small controller is a single-phase converter attached to transmission lines by a single-turn transformer. The converters are hanging on the line so that no costly high-voltage isolation is required. The single-turn transformer uses the secondary winding, inserting the line directly. Each D-FACTS the line and controlled remotely communication (see Fig. 7).

The structure of the D-FACTS unit is shown in Fig. 7. The D-FACTS unit is a single-phase device floating on lines, between phases are avoided. The any transmission-voltage level, supporting phase-ground voltage rating of each unit is units are clamped on transmission



transmission line as the controllable impedance into module is self-powered from by wireless or power-line

FACTS results in low cost FACTS units are single-high-voltage isolations unit can easily be applied at be- cause it does not require isolation. The power and relatively small. Further, the lines, and therefore, no

Fig. 7. D-FACTS unit configuration [7].

land is required. The redundancy of the D-FACTS provides an uninterrupted operation during a single module failure, thereby giving a much higher reliability than other FACTS devices.

C. DPFC Advantages

The DPFC can be considered as a UPFC that employs the D FACTS concept and the concept of exchanging power through harmonic. Therefore, the DPFC inherits all the advantages of the UPFC and the D-FACTS, which are as follows.

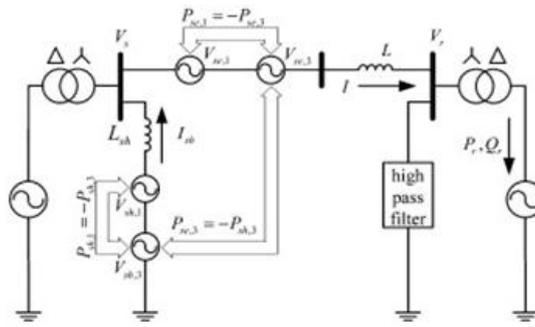
- 1) High control capability. The DPFC can simultaneously control all the parameters of the power system: the line impedance, the transmission angle, and the bus voltage. The elimination of the common dc link enables separated installation of the DPFC converters. The shunt and series converters can be placed at the most effectively location. Due to the high control capability, the DPFC can also be used to improve the power quality and system stability, such as low-frequency power oscillation damping, voltage sag restoration, or balancing asymmetry.
- 2) High reliability. The redundancy of the series converter gives an improved reliability. In addition, the shunt and series converters are independent and the failure at one place will not influence the other converters. When a failure occurs in the series converter, the converter will be short-circuited by bypass protection, thereby having little influence to the network. In the case of the shunt converter failure, the shunt converter will trip and the series converter will stop providing active compensation and will act as the D-FACTS controller.
- 3) Low cost. There is no phase-to-phase voltage isolation required by the series converter. Also, the power rating of each converter is small and can be easily produced in series production lines.

However, as the DPFC injects extra current at the third-harmonic frequency into the transmission line, additional losses in the transmission line and transformer should be aware of.

III. ANALYSIS OF THE DPFC

In this section, the steady-state behavior of the DPFC is analyzed, and the control capability of the DPFC is expressed in the parameters of the network and the DPFC.

To simplify the by controllable voltage Since each converter generates frequencies, it is represented controllable voltage sources, and the other at the third-that the converters and the total active power generated sources will be zero. The simplified as one large is equal to the sum of the shown in Fig. 8.



DPFC, the converters are replaced sources in series with impedance. the voltage at two different by two series-connected one at the fundamental frequency harmonic frequency. Assuming transmission line are lossless, the by the two frequency voltage multiple series converters are converter with the voltage, which voltages for all series converter, as

In Fig. 8, the DPFC is placed in a two-bus system with the sending-end and the receiving-end voltages Vs and Vr , The

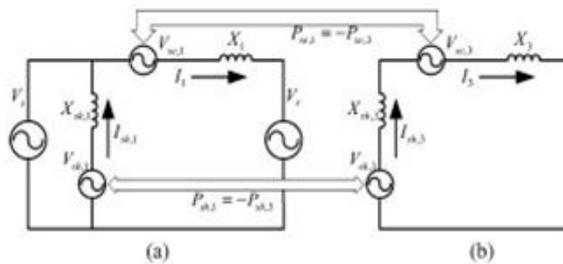


Fig. 8. DPFC simplified

transmission line is the line current I. The series converters is Vse,1 the third-harmonic converter is connected to the sending bus through the inductor Lsh and generates the voltage Vsh,1 and Vsh,3 ; the current is Ish . The active and reactive power flow at the receiving

This representation consists harmonic frequency superposition theorem, the simplified by being split into frequencies. The two circuits the link between these circuits converter, as shown in Fig. 9.

representation.

represented by an inductance L with voltage injected by all the DPFC and Vse,3 at the fundamental and frequency, respectively. The shunt converter is connected to the sending bus through the inductor Lsh and generates the voltage Vsh,1 and Vsh,3 ; the current is Ish . The active and reactive end is Pr and Or, respectively.

of both the fundamental and third-components. Based on the circuit in Fig. 8 can be further two circuits at different are isolated from each other, and is the active power balance of each

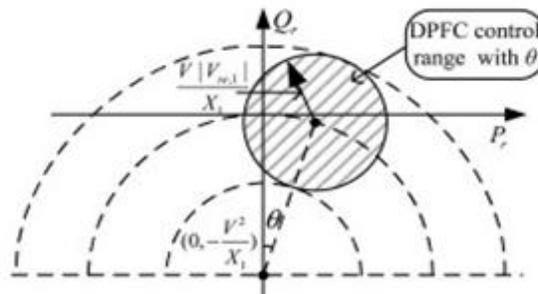


Fig. 9. DPFC equivalent circuit. (a) Fundamental frequency. (b) Third-harmonic frequency.

Fig. 10. DPFC active and reactive power control range with the transmission angle θ .

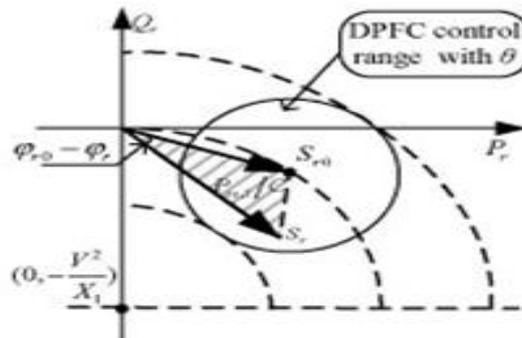
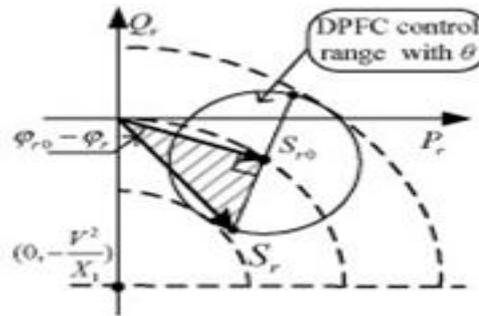
The power-flow control is achieved by the active power at the receiving end. The fundamental frequency active and reactive power

where P_{r0} , Q_{r0} , and θ are the active, reactive power flow, of the uncompensated system, impedance at fundamental voltage magnitude at both ends. In the power flow without the DPFC, the flow can be expressed as follows:

are the active, reactive power flow, of the uncompensated system, impedance at fundamental voltage magnitude at both ends. In the power flow without the DPFC, the flow can be expressed as follows:

To ensure the series rotatable voltage, an active and fundamental frequency is provided by the series converter supplied by the shunt converter. This active power requirement is given by

where ϕ_{r0} is the power angle at the receiving end of the un-compensated system, which equals $\tan^{-1}(P_{r0}/Q_{r0})$ and ϕ_r is the power angle at receiving end with the DPFC compensation. The line impedance X_1 and the voltage magnitude $|V_r|$ are constant; therefore, the required active power is proportional to $|S_r - S_{r0}| \sin(\phi_{r0} - \phi_r)$, which is two times the area of the triangle that is formed by the two vectors S_{r0} and S_r . Fig. 11 illustrates the relationship between $P_{se,1}$ and the power flow at the receiving end at a certain power angle θ .



capability of the DPFC can be illustrated by the active power P_r and reactive power Q_r received. Because the DPFC circuit at the receiving end behaves the same as the UPFC, the flow can be expressed as follows:

are the active, reactive power flow, of the uncompensated system, impedance at fundamental voltage magnitude at both ends. In the power flow without the DPFC, the flow can be expressed as follows:

are the active, reactive power flow, of the uncompensated system, impedance at fundamental voltage magnitude at both ends. In the power flow without the DPFC, the flow can be expressed as follows:

This active power requirement is given by

$$(P_r - P_{r0})^2 + (Q_r - Q_{r0})^2 = \frac{|V_{se,1}|^2}{X_1^2} \quad (2)$$

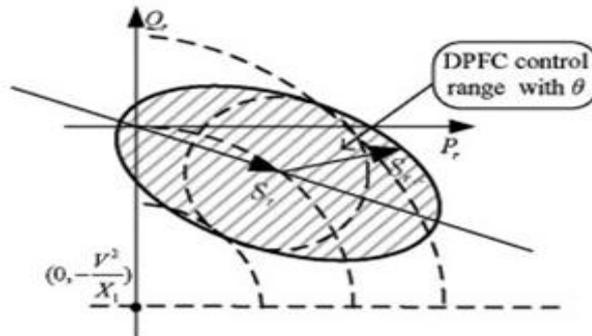
$$P_{se,1} = \text{Re}(V_{se,1} I_r) = \frac{X_1}{|V_r|^2} |S_r - S_{r0}| \sin(\phi_{r0} - \phi_r) \quad (3)$$

$$P_{se,1} = CA(S_r, \phi_{r0}, \phi_r) \quad (4)$$

Fig. 11. Relationship between $P_{se,1}$ and the power flow at the receiving end.

Fig. 12. Maximum active converters.

Consequently, the required converter can be written as



power requirement of the series active power by the series follows:

where the coefficient $C = 2X_1 / |V_r|^2$ and $A(0, r_0, r)$ is the area of the triangle $(0, S_{r0}, S_r)$. The angle difference $\phi_0 - \phi_r$ can be positive or negative, and the sign gives the direction of the active power through the DPFC series converters. The positive sign means that the DPFC series converters generate active power at the fundamental frequency and vice versa. The active power requirement varies with the controlled power flow, and the active power requirement has its maximum when the vector $S_r - S_{r0}$ is perpendicular to the vector S_{r0} , as shown in Fig. 12.

According to Fig. 12, the relationship between the powerflow control range and the maximum active power requirement can be represented by

where $|S_{r,c}|$ is the control range of the DPFC.

Each converter in the DPFC generates two frequency voltages at the same time. Accordingly, the voltage rating of the each converter should be the sum of the maximum voltage of the two frequencies component

$$P_{se,1,max} = \frac{|X_1 S_{r0}|}{|V_r|^2} |S_{r,c}| \quad (5)$$

$$V_{se,max} = |V_{se,1,max}| + |V_{se,3,max}| \quad (6)$$

$$|S_{r,c}| = \frac{|V_r V_{se,1,max}|}{X_1} \quad (7)$$

Fig. 13. DPFC power-flow

control range.

During the operation, of the se-ries converter varies fundamental frequency. When series voltage at the third-smaller than $|V_{se,3,max}|$. This potential voltage that is between $V_{se,3}$ and $|V_{se,3,max}|$ can be used to control the power flow at the fundamental frequency, thereby increasing the power-flow control region of the DPFC. When $S_{r,c}$ is perpendicular to

$$|S_{r,c}| = \frac{|V_r| (|V_{se,1,max}| + |V_{se,3,max}|)}{X_1} \quad (8)$$

the active power requirement with the voltage injected at the the requirement is low, the harmonic frequency will be

the uncompensated power S_r 0 maximum active power, and region is given by

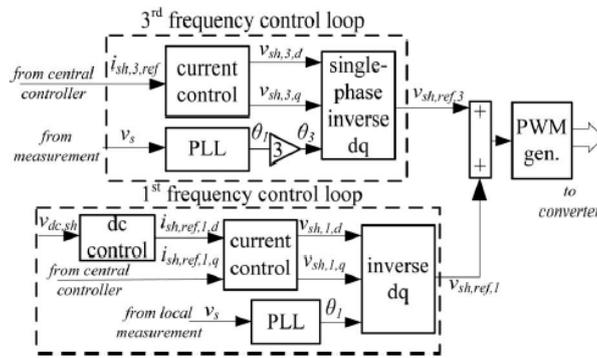
If $S_{r,c}$ is in the same converters only provide the boundary of the DPFC control

It shows that the control extended to a shape that is in Fig. 13.

To obtain the same the rating of the DPFC frequency should be the same

Because the voltages and currents at the third-harmonic frequency have to be added, the rating of the DPFC converter is slightly larger than the UPFC. The increased rating is related with the active power exchanged at the third-harmonic frequency. For a transmission line, the line impedance $|X_1|$ is normally around 0.05 p.u. (per unit). Assuming the bus voltages $|V|$ and uncompensated power flow $|S_r|$ is 1 p.u., and then, from (7), we can see that to control 1-p.u. power flow, the exchanged active power is around 0.05 p.u.

Even with this extra voltage and current at the third-harmonic frequency, the cost of the DPFC is still much lower than the UPFC, for the following reasons: 1) the UPFC converter handles the line-to-line voltage isolation that is much larger than voltage injected by the series converter; 2) no land requirement for the series converter; and 3) the active and passive components for the DPFC converter are low-voltage components (less than 1 kV and 60 A), which is much cheaper than the high-voltage components in the UPFC.



the series converters require the radius of the DPFC control

line as $S_r = 0$, the series reactive compensation and the region will extend to region of the DPFC can be similar as an ellipse, as shown

control capability as the UPFC, converter at the fundamental as the one for the UPFC.

Fig. 14. DPFC control block diagram.

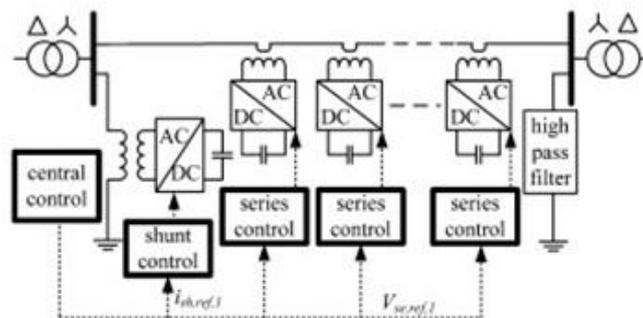


Fig. 15. Block diagram of

IV. DPFC CONTROL

To control the multiple three types of controllers; shunt control, and series

The shunt and

controllers and are responsible for maintaining their own converters' parameters. The central control takes account of the DPFC functions at the power system level. The function of each controller is listed next.

A. Central Control

The central control generates the reference signals for both the shunt and series converters of the DPFC. It is focused on the DPFC tasks at the power-system level, such as power-flow control, low-frequency power oscillation damping,

the series converter control.

converters, DPFC consists of they are central controller, control, as shown in Fig. 14. series control are local

and balancing of asymmetrical components. According to the system requirement, the central control gives corresponding voltage reference signals for the series converters and reactive current signal for the shunt converter. All the reference signals generated by the central control are at the fundamental frequency.

B. Series Control

Each series converter has its own series control. The controller is used to maintain the capacitor dc voltage of its own converter by using the third-harmonic frequency components and to generate series voltage at the fundamental frequency that is prescribed by the central control.

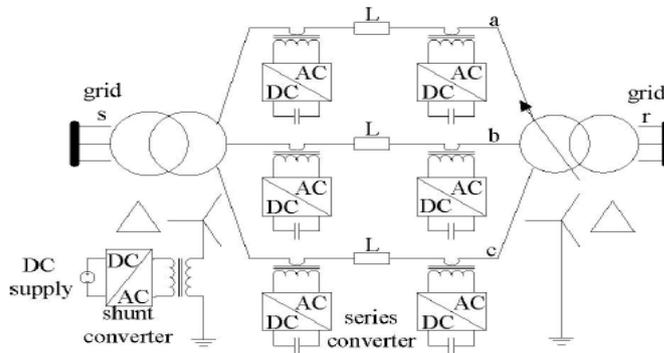
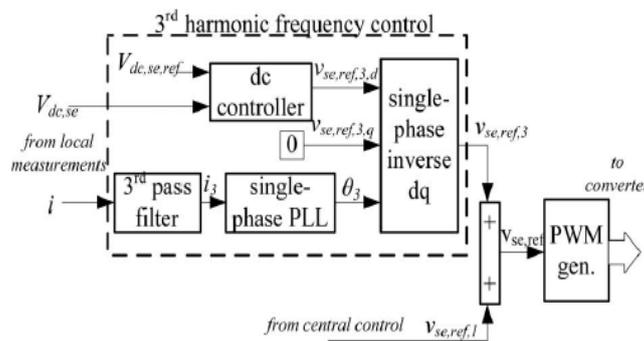


Fig. 16. Block diagram

of the shunt converter control.

The third-harmonic frequency control is the major control loop with the DPFC series converter control. The principle of the vector control is used here for the dc-voltage control. The third-harmonic current through the line is selected as the rotation reference frame for the single-phase park transformation. Because it is easy to be captured by the phase-locked loop (PLL) in the series converter. As the line current contains two frequency components, a third high-pass filter is needed to reduce the fundamental current. The d-component of the third-harmonic voltage is the parameter that is used to control the dc voltage, and its reference signal is generated by the dc-voltage control loop. To minimize the reactive power that is caused by the third harmonic, the series converter is controlled as a resistance at the third-harmonic frequency. The q-component of the third harmonic voltage is kept zero during the operation.

As the series converter is single phase, there will be voltage ripple at the dc side of each converter. The frequency of the ripple depends on the frequency of the current that flows through the converter. As the current contains the fundamental and third-harmonic capacitor voltage will Hz frequency component to reduce this ripple. One the single-phase converter to reduce the flows into the converter. capacitor with a larger



frequency component, the dc-contain 100-, 200-, and 300-. There are two possible ways is to increase the turn ratio of transformer of the series magnitude of the current that The other way is to use the dc capacitance.

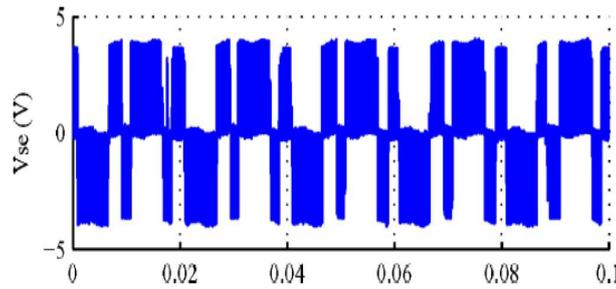
C. Shunt Control

The block diagram of the shown in Fig. 16.

The objective of the shunt control is to inject a constant third harmonic current into the line to provide active power for the series converters. The third-harmonic current is locked with the bus voltage at the fundamental frequency. A PLL is used to capture the bus-voltage frequency, and the output phase signal of the PLL is multiplied by three to create a virtual rotation reference frame for the third-harmonic component. The shunt converter's fundamental frequency control aims to inject a controllable reactive current to grid and to keep the capacitor dc voltage at a constant level. The control for the fundamental frequency components consists of two cascaded controllers. The current control is the inner control loop, which is to modulate the shunt current at the fundamental frequency. The q-component of the refer ence signal of the shunt converter is obtained from the central controller, and d-component is generated by the dc control.

shunt converter control is

Fig. 17. DPFC experimental



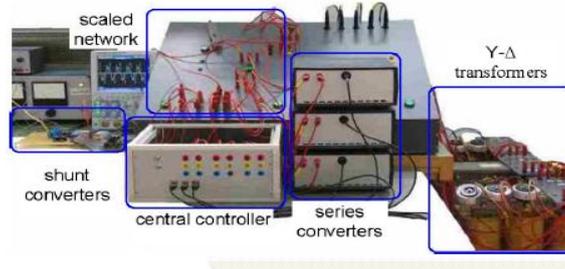
setup circuit.

Fig. 18. DPFC experimental setup.

V. LABORATORY RESULTS

An experimental setup has been built to verify the principle and control of the DPFC. One shunt converter and six single-phase series converters are built and tested in a scaled network, as shown in Fig. 17. Two isolated buses with phase difference are connected by the line. Within the experimental setup, the shunt converter is a single-phase inverter that is connected between the neutral point of the Y- Δ transformer and the ground. The inverter is powered by a constant dc-voltage source. The specifications of the DPFC experimental setup are listed in the Appendix (see Table I).

Within the setup, multiple series converters are controlled by a central controller. The central controller gives the reference voltage signals for voltages and currents within oscilloscope and processed MATLAB. The photograph setup is illustrated in Fig. 18.



To verify the DPFC demonstrated: the DPFC step response. In steady controlled to insert a voltage component, which is $= -0.1$ V. Figs. 19–21 show one operation point of the DPFC setup. For clarity, only the waveforms in one phase are shown. The voltage injected by the series converter, the current through the line, and the voltage and current at the Δ side of the transformer are illustrated.

The the setup are measured by an in computer by using the of the DPFC experimental principle, two situations are behavior in steady state and the state, the series converter is vector with both d- and q- $V_{se,d,ref} = 0.3$ V and $V_{se,q,ref}$

Fig. 19. DPFC operation in steady state: line current.

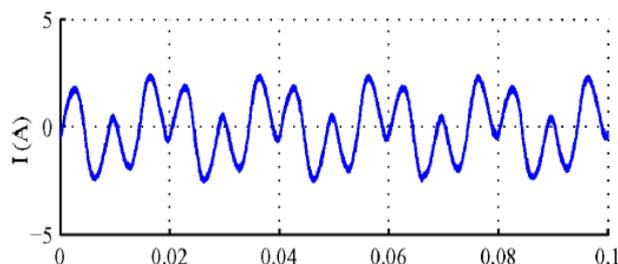


Fig. 20. DPFC operation in

voltage.

steady state: series converter

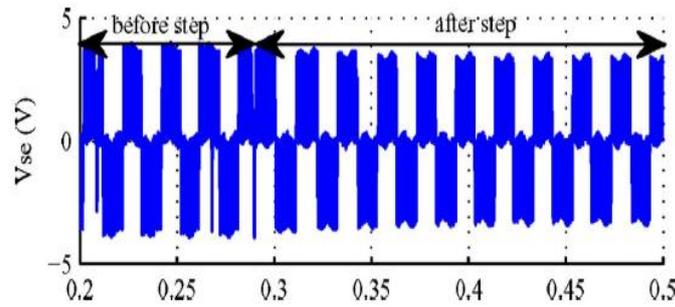


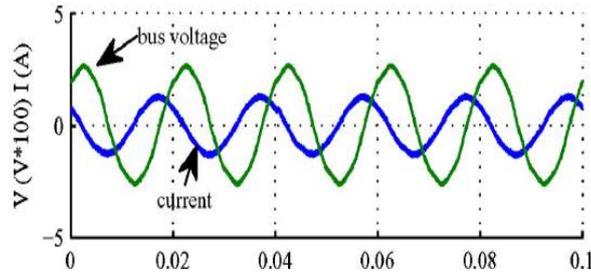
Fig. 21. DPFC operation

in steady state: bus voltage

the transformer.

and current at the of

The constant third-harmonic current is injected by the shunt converter evenly and is superimposed on the fundamental current in Fig. 19. The voltage injected contains two frequency components. The amplitude of the waveform represents the dc voltage well maintained by the third-harmonic filtering by the Y-Δ transformers. There is no third-harmonic current or voltage leaking to the Δ side of the transformer.



harmonic current injected by the series converter also disperses to the three phases fundamental current, as shown by the series converter also components in Fig. 20. The pulsewidthmodulated (PWM) capacitor voltage, which is harmonic component in steady state

The DPFC controls the power flow through transmission lines by varying the voltage injected by the series converter at the fundamental frequency. Figs. 22–26 illustrate the step response of the experimental setup. A step change of the fundamental reference voltage of the series converter is made, which consists of both active and reactive variations, as shown in Fig. 22.

As shown, the dc voltage of the series converter is stabilized before and after the step change. To verify if the series converter can inject or absorb active and reactive power from the grid at the fundamental frequency, the power is calculated from the measured voltage and current in Figs. 23 and 24. The measured data in one phase are processed in the computer by using MATLAB. To analyze the voltage and current at the fundamental frequency, the measured data that contains harmonic distortion are filtered by a low-pass digital filter with the 50-Hz cutoff frequency. Because of this filter, the calculated voltage and current at the fundamental frequency have a 1.5 cycle delay to the actual values, thereby causing a delay of the measured active and reactive power. Fig. 25 illustrated the active and reactive power injected by the series converter. A comparison is made between the measured power and the calculated power. We can see that the series converters are able to absorb and inject both active and reactive power to the grid at the fundamental frequency.

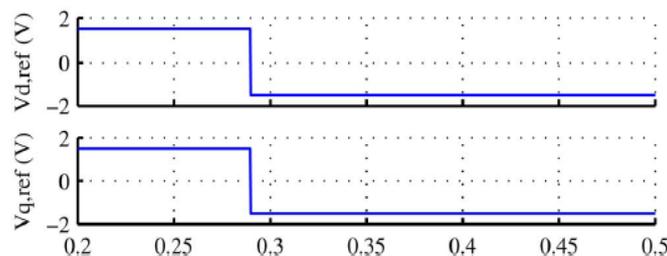


Fig. 22. Reference voltage

for the series converters.

Fig. 23. Step response of the DPFC: series converter voltage.

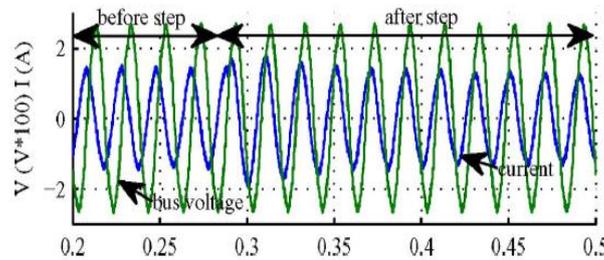


Fig. 24. Step response of the

DPFC: line current.

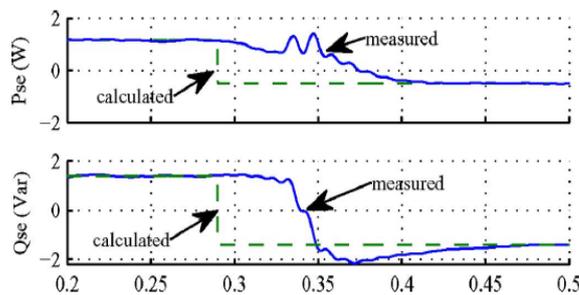
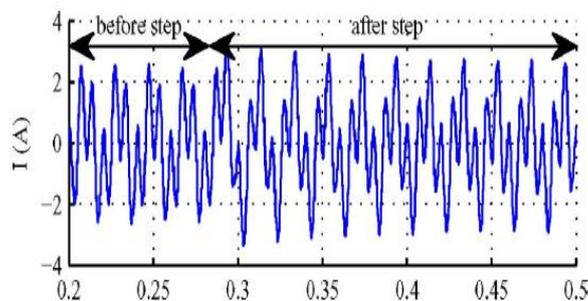


Fig. 25. Step response of the injected by the series converter at the

DPFC: active and reactive power fundamental frequency.

Fig. 26. Step response of the Δ side of the transformer.



DPFC: bus voltage and current at

VI. CONCLUSION

This paper has presented a new concept called DPFC. The DPFC emerges from the UPFC and inherits the control capability of the UPFC, which is the simultaneous adjustment of the line impedance, the transmission angle, and the bus-voltage magnitude. The common dc link between the shunt and series converters, which is used for exchanging active power in the UPFC, is eliminated. This power is now transmitted through the transmission line at the third-harmonic frequency. The series converter of the DPFC employs the D-FACTS concept, which uses multiple small single-phase converters instead of one large-size converter. The reliability of the DPFC is greatly increased because of the redundancy of the series converters. The total cost of the DPFC is also much lower than the UPFC, because no high-voltage isolation is required at the series-converter part and the rating of the components of is low. The DPFC concept has been verified by an experimental setup. It is proved that the shunt and series converters in the DPFC can exchange active power at the third-harmonic frequency, and the series converters are able to inject controllable active and reactive power at the fundamental frequency.

TABLE I
 SPECIFICATION OF THE DPFC EXPERIMENTAL SETUP

Symbol	Description	Value	Unit
V_s	nominal voltage of grid <i>s</i>	220	V
V_r	nominal voltage of grid <i>r</i>	220	V
θ	transmission angle between grid <i>s</i> and <i>r</i>	1	°
L	line inductance	6	mH
$V_{sh,max}$	shunt converter maximum ac voltage	50	V
$I_{sh,max}$	shunt converter maximum ac current	9	A
$V_{sh,dc}$	shunt converter dc source supply	20	V
$I_{sh,ref,3}$	reference 3 rd harmonic current injected by the shunt converter	3	A
f_{sw}	switching frequency for the shunt and series converter	6	kHz
$V_{se,max}$	maximum ac voltage at line side of the series converter	7	V
$I_{se,max}$	maximum ac current at line side of the series converter	15	A

REFERENCES

[1] Y.-H. Song and A. Johns, Flexible ac Transmission Systems (FACTS) (IEE Power and Energy Series), vol. 30. London, U.K.: Institution of Electrical Engineers, 1999.

[2] N. G. Hingorani and L. Gyugyi, Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems. New York: IEEE Press, 2000.

[3] L. Gyugyi, C. D. Schauder, S. L. Williams, T. R. Rietman, D. R.Torgerson,and A. Edris, “The unified power flow controller: A new approach to power transmission control,” IEEE Trans. Power Del.,vol. 10, no. 2, pp. 1085–1097, Apr. 1995.

[4] A.-A. Edris, “Proposed terms and definitions for flexible ac transmission system (facts),” IEEE Trans. Power Del., vol. 12, no. 4, pp. 1848–1853,Oct. 1997.

[5] K. K. Sen, “Sssc-static synchronous series compensator: Theory, modeling, and application,” IEEE Trans. Power Del., vol. 13, no. 1, pp. 241–246,Jan. 1998.

[6] M. D. Deepak, E. B. William, S. S. Robert, K. Bill, W. G. Randal, T. B. Dale, R. I. Michael, and S. G. Ian, “A distributed static series compensator system for realizing active power flow control on existing power lines,” IEEE Trans. Power Del., vol. 22, no. 1, pp. 642–649, Jan.2007.

[7] D. Divan and H. Johal, “Distributed facts—A new concept for realizing grid power flow control,” in Proc. IEEE 36th Power Electron. Spec. Conf. (PESC), 2005, pp. 8–14.

Group features of small seismic waveforms

Wenlong Liu¹, Yucheng Liu²

Shanghai Earthquake Administration, Shanghai, China, 200062

Department of Mechanical Engineering

University of Louisiana, Lafayette, LA70504, USA

Abstract: This paper demonstrates several group features observed from small seismic waveforms and distinguishes abnormal group features that preceded destructive major shocks from normal ones. Important group features discussed in this paper include direction of seismic wave's first motion, amplitude ratio, half period of the first motion, frequency components of earthquakes, and linearity of seismic waveform. The group features illustrated in this paper have been used as important criteria in earthquake prediction.

Keywords: first motion, amplitude ratio, linearity of seismic waveform, consistency, earthquake prediction

I. Introduction

It was learned from numerous earthquake examples that abnormal group features of seismic waves of small earthquakes are usually signs of strong major earthquakes. Therefore, in earthquake prediction, the features and information carried by small seismic waves need to be thoroughly investigated to determine seismic tendency in certain area. This paper presents reveals the correlations between the major shock and several seismic wave features, including direction of the first motion, amplitude ratio, half period of the first motion, frequency components of foreshocks, and linearity of seismic waveform. The illustrative examples and methods introduced in this paper can be used as significant tools in future earthquake prediction.

II. Consistency of Directions of the First Motion

To determine if strong earthquakes will occur in the future based on the consistency of first motion of P waves of small and micro earthquake sequences that are recorded by the seismic observatories that are near to the epicenter was firstly found and applied by the author when he was analyzing the Huoshan earthquake cluster in Anhui Province. An earthquake cluster occurred in Huoshan County in March of 1973, the largest earthquake was $M_L4.5$ that occurred on March 11th. Two large earthquakes occurred before that one, one of them was a $M_L4.3$ earthquake that just occurred eight minutes before the largest shock and the other one was $M_L3.1$ that occurred on March 7th. After the largest shock, another comparatively strong earthquake occurred on March 12th, whose magnitude was $M_L4.0$. Since then, this earthquake cluster attenuated with ups and downs. At that time, the author worked at Fuziling seismic observatory, which was only about 10km from the epicenter. In analyzing seismograms, it was found that before the $M_L4.5$ earthquake, the vertical direction of the initial motion of P waves was downward most of time, except for the short time span after the $M_L3.1$ earthquake, during which the direction was upward (the initial motion direction at the time that $M_L3.1$ occurred was also downward). However, that direction was upward on March 12th when $M_L4.0$ occurred and since then the direction was in disorder. Meanwhile, several other observations were made by the author. (1) Ts-p of small and micro earthquakes before the $M_L4.5$ earthquake was consistent, which was 1.3 seconds. After the $M_L4.0$ earthquake, Ts-p varied from 1.1 ~ 1.5 seconds, which indicated a diffusion of the epicentral distribution range. (2) Before the $M_L4.5$ earthquake, especially before March 7th, the daily frequency of the small and micro earthquakes unconventionally increased day by day and a notable swarm-equanimity phenomenon accompanied. Such daily frequency ceased increasing after the $M_L4.0$ earthquake. (3) Before the $M_L4.5$ earthquake, especially before March 7th, the magnitude of small earthquakes gradually rose. Based on above observations, the author determined that the $M_L4.5$ earthquake must be the largest shock and there would not be any stronger shock occurs after that, which had been proven right. These phenomena recurred in later earthquakes, and the most famous example is the M7.3 Haicheng earthquake. Haicheng earthquake is a massive earthquake with direct foreshock sequence. As recorded by Shipengyu station (Ts-p was 2.5 seconds), in its foreshock sequence, 79 shocks had identifiable initial motion direction and 78 of them had a downward direction. However, such direction was disorderly after the main shock [1]. Haicheng earthquake was the first and the only successful earthquake prediction in history and based on the prediction, Chinese government was able to successfully evacuate much of the populace and furthest mitigate the disaster caused by that earthquake. An essential key of this successful prediction was correctly identifying its foreshock sequence and the consistency of P wave's initial motion direction provided a primary basis for that. Since then, this method became one of most popular methods in judging if a sequence is a foreshock sequence and whether or not strong earthquakes will occur after that.

Table 1. Distribution of initial motion of P waves of Huoshan earthquake swarm (occurred in March of 1973) recorded by Fuziling station (N is the number of earthquakes whose initial motion direction was identifiable, N_U is the number of earthquakes with upward initial motion, N_L is the number of earthquakes with downward motion and % is its percentage, B means before the main shock, and A means after the main shock)

Date (03/x)	4~6	7			8	9	10	11			12			13	14	15	16	17~3 1
		B	M_L 3.1	A				M_L .3	M_L .5	A	B	M_L .0	A					
N	19	28		23	5	13	4			55	30		13	14	11	7	5	31
N_U	0	2		8	1	0	0			14	13		6	6	3	4	2	15
N_L	19	26	X	15	4	13	4	X	X	41	17	X	7	8	8	3	3	16
%	100	92.9		65.2	80	100	100			74.5	56.7		53.8	57.1	72.7	42.9	60	51.6

Using consistency of P wave's initial motion direction to predict future earthquakes is simple and convenient but in some cases, false prediction could be made if only using that method. For example, in July of 1979, small earthquake activities took place in Huoshan with the maximum magnitude $M_{2.2}$ (occurred at 21:00 on July 8th). During that earthquake swarm, overall 27 earthquakes with identifiable initial motion direction were recorded by Fuziling seismic station and it was found that all the earthquakes had a downward initial motion direction except for the one occurred on July 9th. Moreover, the initial motion directions of all 84 small earthquakes recorded by Huoshan seismic station were downward. Therefore, we cannot determine the occurrence of stronger earthquakes only based on the consistency of the initial motion direction and other seismic indicators need to be synthetically analyzed in order to attain accurate predictions. Meanwhile, not every local station can accurately reflect variation of the initial motion direction of small earthquakes. For example, in the Wencheng earthquake swarms (1.2.2), after the maximum earthquake ($M_L4.6$), a consistency of P wave's initial motion direction of the small earthquakes was found according to the records in nearby stations (Hangzhou, Wenzhou, etc.), which became a strong proof that foreboded occurrence of stronger earthquakes. However, it was proven that these stations failed to detect the initial motion direction because they are not close enough to the fault lines. Actually, turbulence of the initial motion direction was only detected by the Xinpuxiang station, a new seismic station established after the earthquake swarms started (7km from the epicenter), after the $M_L4.6$ earthquake. Finally, for those dual-shock-type earthquakes (the magnitude difference between the first main shock (the strongest earthquake) and the second main shock (the second strongest earthquake) is less than 0.3), two groups of small shocks with different fault plane solutions usually occur between the two main shocks. The initial motion direction of the two groups of small shocks may not be consistent. In that case, even the initial motion direction is in disorder, another strong earthquake (the second main shock in a dual-shock-type earthquake) will still occur after the first main shock. Such mechanism is the same as that of the consistency of focal mechanism solutions before mid-strong earthquakes.

III. Consistency of Amplitude Ratio

Amplitude ratio of an earthquake is the ratio between its P wave's maximum vertical amplitude and its S wave's maximum vertical amplitude. Those amplitudes must be recorded by the same instrument. A number of earthquake examples showed that the amplitude ratio of those small earthquakes which occurred before the largest earthquake in a foreshock sequence or an earthquake swarm was fairly consistent. However, that value highly dispersed after the largest earthquake. Fig. 1 presents an example, which is taken from the Haicheng earthquake [2]. Thus, the consistency of amplitude ratio became a popular and effective criterion in deciding whether or not a larger earthquake will occur in the future.

Compare to the mechanism of consistency of amplitude ratio of the small earthquakes, the mechanism of those mid-strong earthquakes is more complicated because it also depends on the medium characteristics. Feng [3] used two-layer crust model and synthetically studied the influences of focal force systems and media on the amplitude ratio. According to Feng, several conclusions were drawn. (1) Amplitude ratio is proportional to the square of the wave velocity ration. Therefore it is possible to observe a visible decrease in the seismic waves that transmitted through the seismogenic zones within quite a long period of time before the main shock. (2) Generally, direction of hypocentral force system considerably affects the amplitude ration. However, such influence is not noticeable for a vertical fault plane that has tangential dislocation. Thus, in earthquake prediction practice, we often compare the amplitude ratios of neighboring earthquakes to attain an accurate prediction. (3) Variation of the amplitude ratio in near-source regions along with hypocentral distance is very intricacy and the amplitude ratio is seriously affected by the azimuth angle. Thus, it is better to use the data recorded by some distant seismic stations, whose distance from the hypocenter ranges from 70 to 100km.

IV. Half Period of Initial Motion

In April 6th of 1976, a $M_L6.3$ earthquake occurred in Hellinger, Inner Mongolia. From the records of Shenliying station (hypocentral distance is about 30km), it was found that among the small and micro earthquakes with consistent initial motion which occurred from April 4th to 8th, 70% of them had a half period of initial motion about 0.1 seconds and only a few of them had 0.2-second half period. However, after April 8th, the number of earthquakes whose half period was 0.1 seconds gradually decreased and the number of earthquakes with 0.2-second half period increased fast [4]. Such trend was also

observed from the data recorded by other local stations. The author also found this phenomenon in his works. For example, in the $M_L 5.0$ Guzhen earthquake which occurred in March 2nd of 1979, the half period of initial motion of its foreshocks was clearly shorter than that of its aftershocks (according to records of Jiashan station). Guzhen earthquake only had a few fore- and aftershocks. The magnitude of the largest foreshock was $M_L 3.2$ and the magnitude of the largest aftershock was $M_L 3.5$. This feature was used by the author in quickly and correctly forecasting the Wencheng earthquake sequence [5, 6].

If the magnitudes are comparable, the half period of initial motion of the foreshock is less than that of the aftershock. Those periods have to be recorded by same station (neglecting the instrumental interference). Such mechanism can have several explanations. Assume the energy released by earthquake is

$$E = \Delta\sigma \cdot D \cdot S \quad (1)$$

where $\Delta\sigma$ is stress drop, D is average dislocation on rupture surfaces, and S is area of the rupture surface. Since the energies released by the earthquakes with comparable magnitudes are close to each other, the less half period of initial motion of the foreshock indicates a smaller area (S) of rupture surface of the fore shock. This means that the energy per unit surface area released during the foreshocks is larger than that released during the aftershocks. In another word, $\Delta\sigma \cdot D$ of the foreshocks is higher than that of the aftershocks. Assuming that D and friction of rupture surface of the foreshocks and aftershocks are roughly the same, it can be deduced that the foreshocks lead to higher ambient stress. Also, for those strong main shocks with high magnitudes, fragmentation of media close to the epicenters may be caused after the main shocks, the Q value will be lowered and more high frequency waves will be absorbed, the half period of initial motion of the aftershock is therefore increased. In addition, for the foreshocks occur under a high effective stress, the rupture will be transmitted at a high speed, which intensifies high frequency radiation and effects of directivity [7, 8]. Like other prediction methods, counter examples were also observed. For instance, it was report that the $M_L 5.7$ Oroville earthquake occurred in August 1st, 1985, the half period of initial motion of its foreshock was longer than that of its aftershock [9]. This example revealed that during earthquakes, the mechanisms of its focal force system and distribution of medium are very intricate.

V. Frequency Components

It has been well accepted that foreshocks have abundant high-frequency components. As observed by Sadovsky et al. [10, 11], in the large processes preceding large earthquakes ($M = 4.5 \sim 5.8$) in the region of Garm, a change in the frequency spectrum of microshocks took place and the percentage of high-frequency shocks increased. Ishida and Kanamori [12] found that the frequency of the spectral peak was systematically higher for the foreshocks than the events prior to 1949 in the small earthquakes preceding the 1952 Kern County, California, earthquake. A similar trend was also found for the 1971 San Fernando earthquake and micro shocks in the region of Xikeer, Xinjiang [13, 14]. Fig. 1 and Table 2 display four small earthquake sequences in the region of Xikeer, Xinjiang. As shown in Table 3, earthquakes in sequence 1, 2, and 3 that occurred before and after the maximum earthquake were grouped in chronological order; earthquakes in sequence 4 were separated into two groups because there was no maximum earthquake in that sequence. Fig. 2 plots average frequency spectra of the four small earthquake sequences occurred in the region of Xikeer. As shown from that figure, the spectrum component between 10Hz and 60Hz increased by 15% during six months before the largest earthquakes in the first three small earthquake sequences. Such variation was not obvious in the fourth sequence. Dominant frequencies of the first three earthquake sequences were about 10Hz and that of the fourth sequence was lower than 6Hz.

For earthquakes with comparable magnitudes, spectra of the foreshocks include more high-frequency component than those of the aftershocks. This mechanism is basically the same as that that governs the half period of initial motion of the foreshocks and aftershocks discussed before. The shorter half period of initial motion of the foreshocks is the reflection of that mechanism in the time domain and the more high-frequency component in the foreshocks is its reflection in the frequency domain.

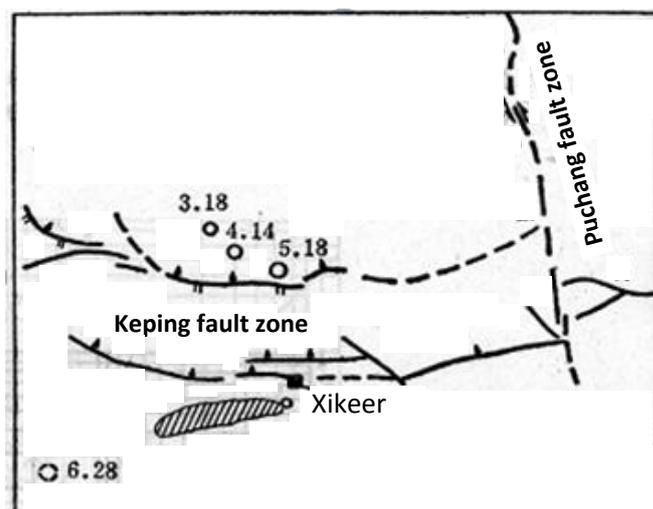


Figure 1. Locations of four small earthquake sequences and local seismic stations in the region of Xikeer

Table 2. Four small earthquake sequences in Xikeer

No.	Duration (in 1972)	Ts-p	M _L	Number of maximum earthquakes					No. of earthquakes	
				Time	Ts-p	M _L	Φ _N	λ _E		h (km)
1	2.11~3.26	2.0~3.0	1.0~2.0	22:37:35.9 on 3.18	2.7	3.7	39°58'	77°15'	25	53
2	3.19~5.23	2.5~3.6	1.0~2.0	19:15:53.9 on 4.14	3.5	4.2	39°57'	77°17'	25	48
3	3.1~5.23	1.5~2.0	0.6~1.0	08:43:42.6 on 5.18	1.7	3.2	39°56'	77°21'	25	45
4	3.29~12.1	4.0~4.6	0.6~1.7	11:05:00 on 6.28	4.5	1.7	Azimuth angle 246°		25	26

Table 3. Grouping of earthquake sequences in Xikeer (the second column lists earthquake groups, third column displays duration of each group, and the fourth column shows number of earthquakes in each group)

1				2				3			4	
A ₁₃	A ₁₂	A ₁₁	B ₁₁	A ₂₂	A ₂₁	B ₂₁	B ₂₂	A ₃₂	A ₃₁	B ₃₁	C ₄₁	C ₄₂
2.11~2.13	2.14~2.29	3.5~3.17	3.19~3.26	3.19~3.31	4.1~4.14	4.14~4.20	4.21~5.23	3.1~4.7	5.2~5.18	5.18~5.23	3.29~4.30	6.28~12.1
11	11	11	9	9	10	11	17	19	15	10	13	13

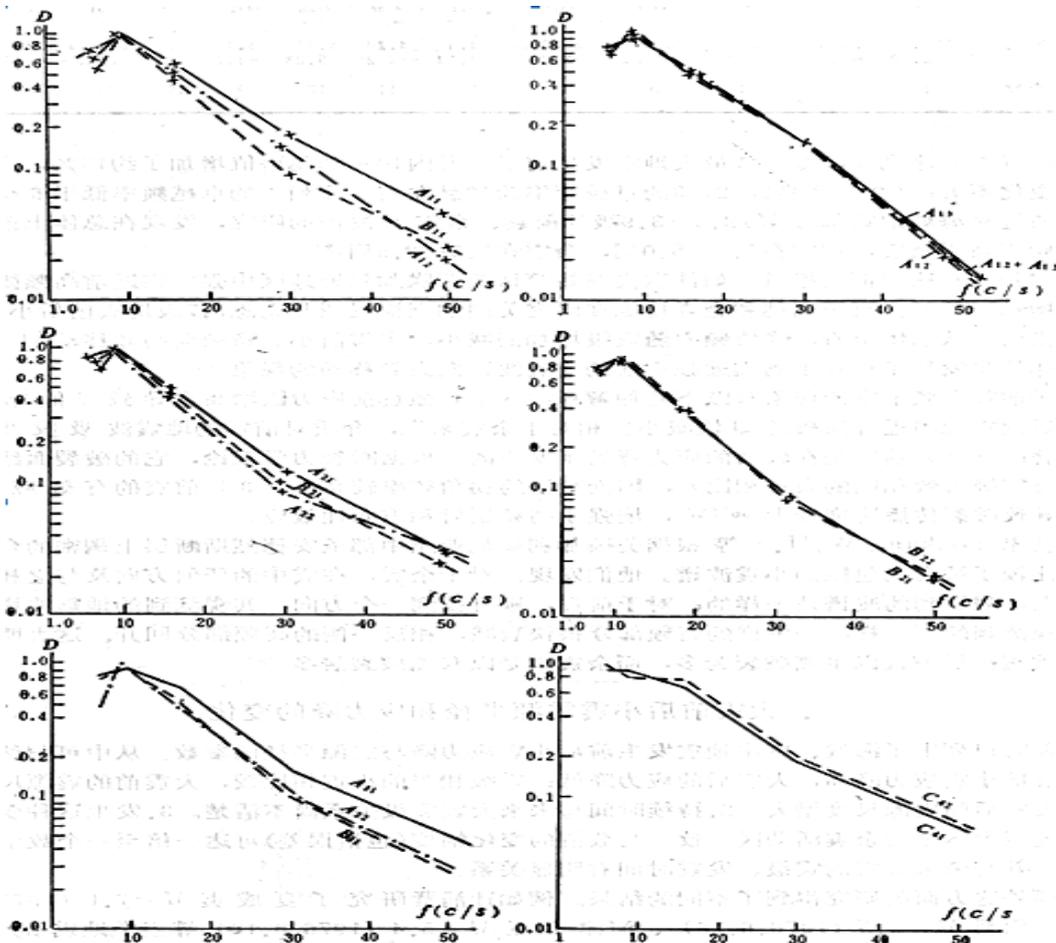


Figure 2. Average frequency spectra of four small earthquake sequences in Xikeer

VI. Linearity of Seismic Waveform

Feng et al. [15] presented temporal and spatial linearity of seismic waveforms from the perspective of system science. Based on Feng's definition, a certain number of time t_1, t_2, \dots, t_n at which the amplitude of displacement or velocity reaches its peak, trough, or zero point are recorded since the P-wave or S-wave initial-motion until one or two wave groups end or n reaches a certain value N . After that, the t_i - i relationship can be drawn in a coordinate system whose ordinate is t_i and abscissa is the index number i , and the linear correlation coefficient r can be calculated using least square method as:

$$r = \frac{S_{ti}}{\sqrt{S_{tt} \cdot S_{ii}}} \quad (2)$$

where

$$\left. \begin{aligned} S_{ii} &= \sum_{i=1}^n \left(t_i - \bar{t} \right) \left(i - \frac{n+1}{2} \right) \\ S_{tt} &= \sum_{i=1}^n (t_i - \bar{t})^2 \\ S_{ii} &= \sum_{i=1}^n \left(i - \frac{n+1}{2} \right)^2 \\ \bar{t} &= \frac{1}{n} \sum_{i=1}^n t_i \end{aligned} \right\} \quad (3)$$

Here r is the temporal linearity of seismic waveforms, and the more the waveforms deviate from the periodic functions, the smaller r is. Slope of the t_i - i curve represents the average half period of oscillation of the initial portion of P waves, $(1/2)T$, which can also be calculated using the least square method.

The spatial linearity can be calculated using following method. Considering the seismic data x_{ij} recorded by a certain observatory as three-dimension random variables, where $i = 1, 2, 3$ and j is sequence number of the sampling points which starts from 1 to the total number of sampling points, N . The expectations and covariance of x_{ij} are

$$E_i = \frac{1}{N} \sum_{j=1}^N x_{ij} \quad (4)$$

$$C_{ik} = \frac{1}{N} \sum_{j=1}^N (x_{ij} - E_i)(x_{kj} - E_k) \quad (5)$$

The covariance matrix V can be constructed based on C_{ik} as $V = \{C_{ik}\}$, $i, k = 1, 2, 3$. From its characteristic equation $(V - \lambda I)n = 0$, its three eigenvalues (from high to low) $\lambda_1, \lambda_2, \lambda_3$ and their corresponding normal eigenvectors n_1, n_2, n_3 can be determined. n_1 and n_3 are incident directions of P and S waves, respectively and the spatial linearity of the P wave can be defined as:

$$\alpha_1 = 1 - \lambda_2/\lambda_1 \text{ or } \alpha_2 = 1 - \lambda_3/\lambda_1 \quad (6)$$

α_1 and α_2 represent the deviation between actual seismic ray path and its initial ray path. The more highly the paths deviate from each other, the smaller α_1 and α_2 are. Thus, α_1 and α_2 reflect homogeneity of the propagation medium, the higher α_1 and α_2 are, the more homogeneous this medium is, and $\alpha_1 = \alpha_2 = 1$ for the homogeneous medium.

Based on analyzing of several medium and strong earthquakes, it was found that the temporal linearity r , spatial linearity α_1 and α_2 , and the half period $(1/2)T$ of the P wave abnormally dropped about one year or longer before the main shock. It was also observed that these anomalies were either intensified or recovered about half years before the main shock. The time span between the "turnover" of the anomalies and the main shock is dependent on the magnitude, epicentral distance, and azimuth angle of the epicenter.

Fig. 3 shows the variations of r, α_1, α_2 , and $(1/2)T$ of P waves of the earthquakes occurred from July of 1988 to October of 1989 in Datong, Shanxi. The displayed curves were plotted based on digital data recorded by Baijiatuan station in Beijing, whose epicentral distance is about 200km. In Fig. 3, time window $\Delta\tau = 4$ seconds, sampling interval is 0.025 seconds and total sampling points N is 2000. The entire duration of the abnormal linearity (ΔT) and the time span between the "turnover" of the anomalies and the main shock (ΔT_1) of several earthquakes are listed in Table 4.

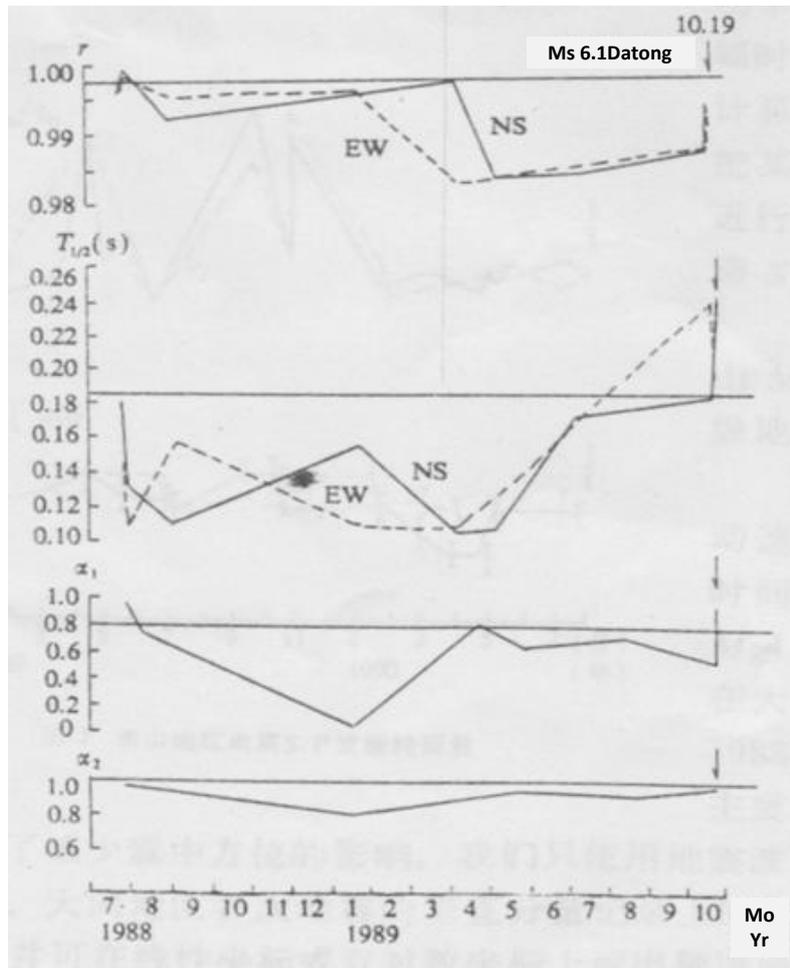


Figure 3. r , α_1 , α_2 , and $(1/2)T$ of seismic P waves in Datong, Shanxi

Table 4. Duration of the abnormal linearity of waveforms of several earthquakes (ΔT and ΔT_1)

Earthquake	Ms	Recorded by	Epicentral distance	ΔT (months)	ΔT_1 (months)
1989.10.0 in Datong	6.1	Baijiatuan	180-200	14	6
1990.7.23 in Tangshan	4.5	Baijiatuan	150	6	3
1990.2.10 in Changshu	5.1	Sheshan	~80	6	1.5-2
1990.4.26 in Gonghe	6.9	Zhang	~250	>16	6
1975.2.3 in Haicheng	7.3	Yingkou	~10	>20	8
		Dandong	~170	>20	2

VII. Conclusions

This paper describes the correlation between precursory abnormal features of small earthquakes' seismic wave and the impending main shocks. From the examples demonstrated in this paper as well as the author's experience, it is concluded that: (1) the direction of the initial motion of P waves becomes consistent before the main shock and stays in disorder during other time; (2) similarly, the amplitude ratio of small earthquakes would be fairly consistent before the strong main shock; (3) small earthquakes whose seismic waves have rich high-frequency components usually foretell impending strong earthquakes; (4) the temporal linearity, spatial linearity, and the half period of the initial motion of P wave might abnormally drop about one year or longer before the main shock.

The criteria and methods presented in this paper are useful in correctly predicting the future earthquakes and identifying the maximum earthquake from an earthquake swarm. It should also be indicated that significant error may be made by singly using one criterion or feature to predict seismic tendency. Accurate earthquake prediction should be made based on synthetically judging and analyzing the variation of different seismic wave features.

Reference

- [1] K.-T. Wu, M.-S.Yue, H.-Y.Wu, X.-L.Cao, H.-T.Chen, W.-Q.Huang, K.-Y.Tian, S.-D. Lu, "Characteristics of Haichen Earthquake Sequence", *ACTA GeophysicaSinica*, 19(2), 1976.
- [2] Y. Jin, Y. Zhao, Y. Chen, J.-Q.Yan, Y.-R.Zhuo, "A Characteristic of Hypocentral Dislocation of Foreshocks of Haicheng Earthquake, Liaoning", *ACTA GeophysicaSinica*, 19(3), 1976.
- [3] D.-Y. Feng, "Abnormal Amplitude Ratio of S and P Waves of Nearby Earthquake Activities and Their Prediction", *ACTA GeophysicaSinica*, 17(3), 1974.
- [4] "Seismic Activity Characteristics of Hellinger Earthquake", Earthquake Team of Inner Mongolia Autonomous Region, *Earthquake*, 5, 1976.
- [5] W.-L. Liu and H.-Y. Yu, "Spot Forecast of the M_L 4.6 Earthquake Swarm of Shanxi Reservoir, Zhejiang Province occurred on Feb. 9, 2006", *South China Journal of Seismology*, 26(3), 2006, 34-44.
- [6] W.-L. Liu and Y.-C.Liu, "Application of Seismic Wave Method in Early Estimation of Wencheng Earthquake", *International Journal of Natural Sciences and Engineering*, 1(2), 2009, 121-128.
- [7] Y.-Z. Lu, J.-W.Shen, "Study of Seismic Rupture Process", *ACTA GeophysicaSinica*, 23(2), 1980.
- [8] T. Johnson, F.T. Wu, C.H. Scholz, "Source Parameters for Stick-Slip and for Earthquakes", *Science*, 179(4070), 1973, 278-280.
- [9] A.G. Lindh, C.E. Mantis, "Premonitory Amplitude Changes", Summaries of Technical Reports, VI, 1976, 184-188.
- [10] M.A. Sadovalsky, I.L. Nersesov, S.K. Nigmatullaev, L.A. Latynina, A.A. Lukk, A.N. Semenov, I.G. Simbireva, V.I. Ulomov, "The Processes Preceding Strong Earthquakes in Some Regions of Middle Asia", *Tectonophysics*, 14(3-4), 1972, 295-307.
- [11] M.A. Sadovalsky, I.L. Nersesov, "Forecasts of Earthquakes on the Basis of Complex Geophysical Features", *Tectonophysics*, 23(3), 1974, 247-255.
- [12] M. Ishida, H. Kanamori, "Temporal Variation of Seismicity and Spectrum of Small Earthquakes Preceding the 1952 Kern County, California, Earthquake", *Bulletin of the Seismological Society of America*, 70(2), 1980, 509-527.
- [13] C.-Z. Zhu, R.-B.Shi, S.-L.Luo, Z.-W.An, "Preliminary Study of Spectra of Microshocks in the Region of Xikeer, Xinjiang", *ACTA GeophysicaSinica*, 18(4), 1975.
- [14] C.-Z. Zhu, C.-H.Fu, S.-L.Luo, "Focal Parameters of Micro Earthquakes Occurred before and after Tangshan M_L 7.8 Earthquake", *ACTA GeophysicaSinica*, 20(4), 1977.
- [15] D.-Y. Feng, G.-Y.Wu, H.-R.Chen, R.-Z.Guo, X.-J.Yu, W.-G. Ding, "Application of Indicator of Dynamic Characteristics of Seismic Waves in Short-Term Earthquake Prediction", *Earthquake*, 1(1), 1994.

“Experimental study on the properties of concrete made with alternate construction materials”

Prof. R. Sathish Kumar

Sr. Associate Professor

National Institute of Construction Management and Research, Hyderabad

Abstract: Rapid increase in construction activities leads to active shortage of conventional construction materials due to various regions. Concrete is most widely used construction material. Cement, sand & granite stone are the constituents of the concrete. Researches were researching for cheaper materials that can be used as substitute for these materials. In this context an experimental study was carried out to find the suitability of the alternate construction materials such as, rice husk ash, sawdust, recycled aggregate and brickbats as a partial replacement for cement and conventional aggregates. For this concrete cubes of size 150mm x150mm were casted with various alternate construction materials in different mix proportion and with different water cement ratios. Their density, workability and compressive strengths were determined and a comparative analysis was done in terms of their physical properties and also cost savings. Test results indicated that the compressive strength of the OPC/RHA concrete cube blocks increases with age of curing and decreases as the percentage of RHA content increases. It was also found that the other alternate construction materials like saw dust, recycled aggregates and brick bats can be effectively used as a partial replacement for cement and conventional aggregates. The results showed that the compressive strength, of recycled aggregate are on average 70% to 80% of the natural aggregate concrete and the compressive strength of brick bat concrete and saw dust concrete was found to be in the range of 30-35% and 8-10% respectively. The compressive strength of rice husk ash concrete was found to be in the range of 70-80% of conventional concrete for a replacement of cement up to 20%.

I. Introduction

Concrete is most widely used construction material today. Concrete has attained the status of a major building material in all the branches of modern construction. It is difficult to point out another material of construction which is as variable as concrete. Concrete is the best material of choice where strength, durability, impermeability, fire resistance & absorption resistance are required. Rice husk & saw dust are the waste products which are abundantly available & which can be used as a substitute for white cement.

I.1 Rice husk ash (RHA) concrete

Rice husk is an agro-waste material which is produced in about 100 million of tons. Approximately, 20 Kg of rice husk are obtained from 100 Kg of rice. Rice husks contain organic substances and 20% of inorganic material. Rice husk ash (RHA) is obtained by the combustion of rice

husk. The burning temperature must be within the range of 600 to 800°C. The ash obtained has to be grounded in a ball mill for 30 minutes and its appearance in colour will be grey. The most important property of RHA that determines pozzolanic activity is the amorphous phase content. RHA is a highly reactive pozzolanic material suitable for use in lime-pozzolana mixes and for Portland cement replacement. RHA contains a high amount of silicon dioxide, and its reactivity related to lime depends on a combination of two factors, namely the non-crystalline silica content and its specific surface. Research on producing rice husk ash (RHA) that can be incorporated to concrete and mortars are not recent. In 1973, investigations were done on the effect of pyroprocessing on the pozzolanic reactivity of RHA. Since then, a lot of studies have been developed to improve the mechanical and durability properties of concrete.

I.2 Previous Research Efforts

The following research efforts shed light on the research works on the utilization of rice husk and rice husk ash as a partial replacement material or stabilizing agent in building works. Tests were carried out on some characteristics of rice husk ash/ordinary Portland cement concrete. Test results indicated that the compressive strength for all the mixes containing RHA increases with age up to the 14-day hydration period but decreases to the 28-day hydration period while the conventional concrete increases steadily up to 28-day hydration period [1]. Tests were also carried on the use of rice husk ash in concrete. Test results indicated that the most convenient and economical temperature required for conversion of rice husk into ash is 500°C. Water requirement decreases as the fineness of RHA increases. The higher the percentage of RHA contents, the lower the compressive strengths [3]

I.3 Saw dust concrete

It is sometimes required to make nailing concrete & this may be achieved by using saw dust as an aggregate. Nailing concrete is a material into which nails can be driven & in which they are firmly held. The last stipulation is made because, for instance in some of the lighter light weight concrete nails, although easily driven, fail to hold. The nailing properties are required in some types of roof construction & pre cast unit for houses etc. because of its very large moisture movement, saw dust concrete should not be used in the situation where it is exposed to moistures (4).

I.4 Recycled aggregate concrete

Lots of construction activities are going on in & around the world & lots of demolition of old concrete works are also taking place. This demolished concrete, if it can be recycled & used as recycled aggregate concrete their disposal which is gigantic task can never be problem. Shanker and Ali(5) have studied engineering properties of rock flour and reported that the rock flour can be used as alternative material in place of sand in concrete based on grain size data. Nagaraj and Banu(9) have studied the effect of rock dust and pebble as aggregate in cement and concrete. It has been reported that crushed stone dust can be used to replace the natural sand in concrete. Sahu, et al (2) have reported that sand can be replaced by rock flour up to 40% without affecting strength and workability. Kanakasabai and Rajashekar(10) investigated the potential of ceramic insulator scrap as coarse aggregate in concrete. It has been reported that the crushed ceramic aggregate can be used to produce lightweight concrete, without affecting strength

I.5 Obstacles in Use of recycled aggregate

The acceptability of recycled aggregate is impeded for structural applications due to the technical problems associated with it such as weak interfacial transition zones between cement paste and aggregate, porosity and transverse cracks within demolished concrete, high level of sulphate and chloride contents, impurity, cement remains, poor grading, and large variation in quality (8). Although, it is environmentally & economically beneficial to use RCA in construction, however the current legislation and experience are not adequate to support and encourage recycling of construction & demolished waste in India. Lack of awareness, guidelines, specifications, standards, data base of utilization of RCA in concrete and lack of confidence in engineers, researchers and user agencies is major cause for poor utilization of RCA in construction. (7)

I.6 Brick bat aggregate concrete

Bricks bats one of the types of aggregates used in certain places where natural aggregates are not available or costly. Where ever brick bats aggregates are used the aggregates are made from slightly over burnt bricks. This will be hard & absorb less water.

II. Experimental Investigations

II.1 Introduction

The experimental investigations includes the casting of cube with various alternative construction materials & the tests were conducted to study the various physical properties such as density, slump, 7days & 28 days compressive strength. A total of 168 specimens were cast & tested in the laboratory to evaluate their compressive strength.

II.2 Materials and Methods

II.2.1 Materials

- **Cement**- Cement used in this study is "43 Grade" which is available under the commercial name "Rajashree Cement".
- **Sand** - River sand confirming to zone -2 and with a fineness modulus of 2.4 was used in this study.

- **Rice Husk Ash** -The rice husk ash was obtained from N.K.Enterprises ,Jharsuguda, Orissa, India . Here rice husk was burnt approximately 48 hours under uncontrolled combustion process. The burning temperature was within the range of 600 to 800°C. The ash obtained was grounded in a ball mill for 30 minutes and its appearance in colour was grey.
- **Coarse aggregate** - From near by quarry.
- **Saw dust** - from nearby saw mill
- **Brick bats** –Collected locally and then broken into pieces of 40mm size, mechanically sieved through 4.75mm sieve to remove the finer particles.
- **Recycled aggregates** The recycled aggregate are collected from the source demolished structures. The concrete debris were collected locally from different sources and broken into the pieces of approximately 80 mm size with the help of hammer .The foreign matters were sorted out from the pieces. Further, those pieces were mechanically sieved through sieve of 4.75 mm to remove the finer particles. The recycled coarse aggregates were washed to remove dirt, dust etc. and collected for use in concrete mix.

II.3 Material tests (As per I S 456-2000)

II.3.1 Cement and rice husk ash

- Initial setting time of cement – 95 min
- Final setting time of cement – 420 min
- Specific gravity of cement – 3.12
- Specific gravity of rice husk ash -2.14
- Fineness of cement – 1.35%
- Setting Times-The comparison of setting times of cement and rice husk ash is presented in table -1

The initial and final setting times increases with increase in rice husk ash content. The reaction between cement and water is exothermic leading to liberation of heat and evaporation of moisture and consequently stiffening of the paste. As rice husk ash replaces cement, the rate of reaction reduces, and the quantity of heat liberated also reduces leading to late stiffening of the paste. As the hydration process requires water, greater amount of water was also required for the process to continue.

II.3.2 Chemical analysis of rice husk ash as supplied by the supplier

Table shows the chemical composition of rice husk ash. The total percentage composition was found to be 73.77% This value is within the range of required value of 70% minimum for pozzolonas. The loss of ignition obtained was 17.71% which is slightly more than 12% max required for pozzolonas. It means that rice husk ash contains little unburnt carbon and this reduces the pozzolonic activity of ash.

Content	% composition
Fe ₂ O ₃	0.91%
SiO ₂	65.3%
CaO	1.31%
Al ₂ O ₃	4.4%
MgO	1.85%

Loss of ignition	17.71%
------------------	--------

The specific gravity of rice husk ash was found to be 2.14. The value is well within the range of pulverised fuel ash which is in between 1.9 and 2.4 as reported in (6)

II.3.3 Fine aggregate tests

Specific gravity of sand – 2.61

Fineness modulus of sand – 2.54

Fineness modulus of saw dust – 2.67

Specific gravity of saw dust -2.12

II.3.4 Coarse aggregates tests

II.3.4.1 Conventional aggregate

Specific gravity of coarse aggregate – 2.70

Fineness modulus of coarse aggregate – 7.133

Crushing strength – 16.1%

Specific gravity of brick bats -2.34

II.3.4.2 Recycled Concrete Aggregate

(a) Specific Gravity and Water Absorption

The specific gravity (saturated surface dry condition) of recycled concrete aggregate was found to be 2.28 which is lower as compared to natural aggregates. Since the RCA from demolished concrete consist of crushed stone aggregate with old mortar adhering to it, the water absorption found to be 2.95% which is relatively higher than that of the natural aggregates.

(b) Bulk Density

The rodded & loose bulk density of recycled aggregate is lower than that of natural aggregate. The lower value of loose bulk density of recycled aggregate may be attributed to its higher porosity than that of natural aggregate.

(c) Crushing and Impact Values

The recycled aggregate is relatively weaker than the natural aggregate against mechanical actions. As per IS 2386, the crushing and impact values for concrete wearing surfaces should not exceed 45% and 50% respectively. The crushing & impact values of recycled aggregate satisfy the BIS specifications

II.4.4.3 Super plasticizer

Product - Roff block master

Description – It is an admixture for making concrete blocks and pre cast concrete products- to achieve homogeneous highly workable mix to give high early strength & to reduce breakages.

Typical applications – For manufacture of concrete blocks, products such as pipes, pole, manhole covers, concrete jails etc.

Consumption – 140ml/bag of 50 Kg cement

Application – Dry mix cement & aggregate, add 140ml roff block master in gauging water, mix thoroughly & cast as per standard practice, permits use of leaner mixes.

II.5 Specimens

A total of 168 specimens of size 15cm x 15cm x 15cm were casted with different alternative construction materials with varying mix proportion & water cement ratio is given in table -2

III. Comparison of results

III.1. Rice husk ash concrete with conventional concrete

Nearly 60 specimens were casted with the mix proportion of 1:2:4 with different water cement ratios 0.56, 0.58, 0.60 and 0.62 and with different percentages of Cement and rice husk ash. It was found from that when the rice husk ash percentage was increased from 50% onwards the mix was becoming harsh, so a higher water cement ratio was adopted for the mixes with increased rice husk ash percentages. The comparison of compressive strength of conventional and rice husk ash concrete cubes are given in table -3

III. 2. Saw dust concrete with conventional concrete

Some properties of concrete with sawdust ash (SDA) as a replacement for conventional fine aggregate are investigated. The cube specimens were casted under 2 mix proportions: 1:2:4 and 1:1.5:3. Sawdust and river sand were taken in the ratio 1.5:0.5 for 1:2:4 concrete and 1:0.5 and 1.25:0.25 for 1:1.5:3 concrete. The water cement ratio adopted earlier was found harsh for this concrete. So the mixes are made in a water cement ratio of 0.7, 0.75 and 0.8 respectively. The compressive strength of specimens with replacement levels shown above cured for periods of 7-28 days showed a decreasing strength with higher saw dust content. The comparison of compressive strength of conventional and saw dust concrete cubes are given in table -4

III.3. Recycled aggregate concrete with conventional concrete

Three different mix proportions 1:1.5:3, 1:2:4, and 1:3:6 with water cement ratio 0.5 & 0.6 and 0.7 were made with natural aggregate concrete and recycled aggregate concrete with and without plasticizer. Due to the higher water absorption capacity of RCA as compared to natural aggregate, both the aggregates are maintained at saturated surface dry (SSD) conditions before mixing operations. The ordinary Portland cement of 43 grade and natural river sand were used throughout the casting work. The maximum size of coarse aggregate used was 20 mm in both recycled and natural aggregate concrete. A total of 54 cube specimens were casted. The comparison of compressive strength of conventional and recycled aggregate concrete cubes are given in table -5

III.4 Brickbat concrete with conventional concrete

Here the conventional stone aggregates were replaced with brickbats and the specimens were casted in the ratio 1:2:4:1:1.5:3 and 1:3:6. The comparison of compressive strength of conventional and brick bat aggregate concrete cubes are given in table -6

IV. Cost analysis

A cost comparison was done with concrete made up of alternate construction materials and conventional concrete was done the same was presented in table 7 & 8

IV.1 Analysis of results obtained

IV.1.1 Suitability of Material Used

- The rice husk ash used was found to be pozzolonic in nature. The specific gravity of rice husk ash was found to be 2.14. The setting time of Ordinary Portland cement and rice husk ash paste increases as rice husk ash content increases
- The fineness modulus of saw dust was 2.67 and that of sand was 2.54. This infers that the saw dust contains more coarse particles in comparison to sand. Hence, saw dust may be considered as fine aggregate in concrete.
- The specific gravity of the brick bat aggregates (2.34) is about 0.86 times as that of conventional aggregate (2.7). As the specific gravity of brick bat aggregates is less than that of the conventional aggregates, the concrete produced using the brick bat aggregate will be of low density.
- Recycled aggregate and brick bat aggregates can be partially used to replace conventional coarse aggregates (10% to 20%), without affecting its structural significance

V. Conclusion

- The compressive strength of rice husk ash concrete was found to be in the range of 70-80% of conventional concrete for a replacement of cement up to 20%.
- The study shows that the early strength of rice husk ash concrete was found to be less and the strength increased with age.
- The rice husk ash concrete occupies more volume than cement for the same weight. So the total volume of the rice husk ash concrete increases for a particular weight as compared to conventional concrete which results in economy.
- Due to the lower density of RHA concrete the self weight of structure gets reduced which results in overall savings.
- From the cost analysis it was found that the cost of RHA concrete was less compared to conventional concrete
- Recycled aggregate possesses relatively lower bulk density, crushing and impact values and higher water absorption as compared to natural aggregate.
- The compressive strength of recycled aggregate concrete was found to be in the range of 70 to 80 % of conventional concrete..
- The compressive strength of brick bat concrete was found to be nearly 35 % of conventional concrete... The compressive strength of saw dust concrete was found to be nearly 10 to 15% of conventional concrete. So the concrete made with alternate construction materials like brick bats and saw dust can be used for partition & filling purposes & nailing purposes where the strength is not the criteria.
- Wherever compressive strength is not a criteria, the concrete made with alternate construction materials can always be preferred.

1. Al -Khalaf M.N. and Yousuf H.A, 'Use of rice husk ash in concrete, The international journal of cement composite and light weight concrete', 6(11), p.241-245, 1984
2. A K Sahu, S Kumar and A K Sachin. 'Crushed Stone Waste as Fine Aggregate for Concrete'. The Indian Concrete Journal, January 2003
3. Dashan .I.I.B and Kamang E.E.I., 'Some characteristics of RHA/OPC concrete, A preliminary assessment', Nigerian journal of Construction Technology and Management, 2(1), p.22-28, 1999
4. Felix F. Udoeyo¹ and Philibus U. Dashibil, J. Mat. in Civ. Engrg. Volume 14, Issue 2, pp. 173-176 (March/April 2002)
5. N B Shankar and Md Ali. 'Engineering Properties of Rock Flour'. National conference on Cement and Building Materials from Industrial Waste, July 24-25, 1992, pp 167-172.
6. Neville AM., 'Properties of concrete', John & Wiley and sons, Malasia, 1981 & 1985 composites and light weight concrete, 6(4), p.241-248, 1984.
7. Rahal, K. (2007) "Mechanical Properties of Concrete with Recycled Coarse Aggregate," Building & Environment, Vol. 42, pp 407-415.
8. Sharma, P.C., Nagraj, N. (1999), "Recycled Aggregate Concrete and Its Importance in Indian Conditions"- All India Seminar on Indian Cement Industries : Challenges and Prospects of Cement" Chandrapur (Maharashtra)
9. T S Nagaraj and Z Banu. 'Efficient Utilization of Rock Dust and Pebbles as Aggregate in Portland Cement Concrete'. The Indian Concrete Journal, vol 70, no 1, 1996, pp 1-4
10. VK Sabai and AR Shekaran. 'Ceramic Insulator Scrap as Light Weight Concrete'. National Conference on Cement and Building Materials from Industrial Wastes, July 4-25, 1992, pp 8-15

References

Table -1 Comparison of setting times of cement and rice husk ash under various replacement levels.

RHA replacement of OPC (%)	0	10	20	30	40	50
Initial Setting Time (minutes)	90	172	183	285	324	389
Final Setting Time (minutes)	240	313	490	525	712	790

Table -2 Specimens casted with different alternate construction materials and with different water Cement ratios.

SI No	Types of concrete	Proportions	W/C ratio	Density (Kg/cum)
1.	Conventional concrete	1:2:4	0.4,0.5,0.6	2520
2.	Concrete with rice husk ash	1:2:4(80% cement and 20% rice husk ash)	0.4,0.5,0.6	2418
4.	Recycled concrete without plasticizer	1:1.5:3,1:2:4,1:3:6	0.5,0.6,0.7	2488
6.	Saw dust concrete	1:(1.5+0.5):4	0.7,0.75,0.8	1980
7.	Concrete using brick bat	1:2:4	0.5,0.6,0.7	2215

Table -3 Comparisons of compressive strength of conventional and rha concrete cubes.

SI No	Types of concrete	Mix	Cement : RHA	W/C	Compressive strength in Kg/cm ²		
					3 days	7 days	28 days
1	Conventional concrete	1:2:4	100% :0%	0.56	118.57	166.00	298.06
2	RHA concrete	1:2:4	90%:10%	0.56	98.69	138.57	234.39
3	RHA concrete	1:2:4	80%:20%	0.56	89.69	124.62	197.98
4	RHA concrete	1:2:4	70%:30%	0.56	71.02	104.00	157.90
5	Conventional concrete	1:2:4	100% :0%	0.58	122.3	168.22	276.67
6	RHA concrete	1:2:4	90%:10%	0.58	100.79	144.74	237.15
7	RHA concrete	1:2:4	80%:20%	0.58	89.05	128.93	198.57
8	RHA concrete	1:2:4	70%:30%	0.58	78.50	106.26	158.12
9	Conventional concrete	1:2:4	100% :0%	0.60	101.67	154.15	277.86
10	RHA concrete	1:2:4	90%:10%	0.60	101.12	146.15	239.33
11	RHA concrete	1:2:4	80%:20%	0.60	91.47	132.98	199.76
12	RHA concrete	1:2:4	70%:30%	0.60	79.66	108.59	161.33
13	Conventional concrete	1:2:4	100% :0%	0.62	120.55	160.43	248.33
14	RHA concrete	1:2:4	90%:10%	0.62	84.05	112.79	216.71
15	RHA concrete	1:2:4	80%:20%	0.62	78.28	107.05	198.81
16	RHA concrete	1:2:4	70%:30%	0.62	61.26	85.09	164.62
17	RHA concrete	1:2:4	50%:50%	0.98	49.4	57.31	75.09
18	RHA concrete	1:2:4	40%:60%	1.07	35.57	55.33	59.28
19	RHA concrete	1:2:4	30%:70%	1.20	25.57	39.52	41.5
20	RHA concrete	1:2:4	20%:80%	1.35	15.81	23.71	25.69

Table-4 Compressive strength test results of conventional concrete and saw dust concrete.

SI No	Types of concrete	Prop	W/C	Compressive Strength (7days) Kg/cm ²	Compressive Strength (28days) Kg/cm ²
	Conventional concrete	1:2:4	0.56	166	328
	Saw dust concrete	1:(1.5+0.5):4	0.7	18.89	31.11
	Saw dust concrete	1:(1.5+0.5):4	0.75	19.09	32.36
	Saw dust concrete	1:(1.5+0.5):4	0.8	17.78	30.35
	Saw dust concrete	1:(1+0.5):3	0.7	20.12	42.11
	Saw dust concrete	1:(1+0.5):3	0.75	20.26	43.22
	Saw dust concrete	1:(1+0.5):3	0.8	19.19	39.32
	Saw dust concrete	1:(1.25+0.25):3	0.7	18.11	30.33
	Saw dust concrete	1:(1.25+0.25):3	0.75	18.63	30.44
	Saw dust concrete	1:(1.25+0.25):3	0.8	18.42	30.14

Table-5 Compressive strength test results of conventional concrete and recycled aggregate concrete.

SI No	Types of concrete	Prop.	W/C	Compressive Strength (7days) Kg/cm ²	Compressive Strength (28days) Kg/cm ²
1	Conventional concrete	1:2:4	0.5	166	296
2	Conventional concrete	1:1.5:3	0.5	198	346
3	Conventional concrete	1:3:6	0.5	136	196
4	Recycled concrete without plasticizer	1:2:4	0.5	101.44	177.56
5	Recycled concrete without plasticizer	1:2:4	0.6	104.67	187.78
6	Recycled concrete without plasticizer	1:2:4	0.7	114.44	192.22
7	Recycled concrete without plasticizer	1:1.5:3	0.5	119.23	218.11
8	Recycled concrete without plasticizer	1:1.5:3	0.6	122.89	221.56
9	Recycled concrete without plasticizer	1:1.5:3	0.7	129.28	228.33
10	Recycled concrete without plasticizer	1:3:6	0.5	81.88	118.66
11	Recycled concrete without plasticizer	1:3:6	0.6	87.78	121.22
12	Recycled concrete without plasticizer	1:3:6	0.7	92	128.78
13	Recycled concrete with plasticizer	1:2:4	0.5	134.44	234.56
14	Recycled concrete with plasticizer	1:2:4	0.6	136.67	241.78
15	Recycled concrete with plasticizer	1:1.5:3	0.5	148.23	258.13

16	Recycled concrete with plasticizer	1:1.5:3	0.6	152.78	272.11
17	Recycled concrete with plasticizer	1:3:6	0.5	109.11	152.06
18	Recycled concrete with plasticizer	1:3:6	0.6	111.43	154.11

Table-6 Compressive strength test results of conventional concrete and brick bat concrete.

SI No	Types of concrete	Prop.	W/C	Compressive Strength (7days) Kg/cm ²	Compressive Strength (28days) Kg/cm ²
1	Conventional concrete	1:2:4	0.56	166	328
2	Concrete using brick bat	1:2:4	0.5	58.12	91.11
3	Concrete using brick bat	1:2:4	0.6	62.89	108.89
4	Concrete using brick bat	1:2:4	0.7	59.67	102.89
5	Concrete using brick bat	1:1.5:3	0.5	64.12	122.33
6	Concrete using brick bat	1:1.5:3	0.6	69.67	132.11
7	Concrete using brick bat	1:1.5:3	0.7	66.78	128.44
8	Concrete using brick bat	1:3:6	0.5	42.89	66.67
9	Concrete using brick bat	1:3:6	0.6	43.22	72.22
10	Concrete using brick bat	1:3:6	0.7	41.08	68.34

Table -7 Cost comparisons between rice husk ash concrete specimens and conventional concrete specimens.

SI No	Types of concrete	Cement:RHA	Prop.	Percentage w.r.t conventional concrete
1	Conventional concrete	100% :0%	1:2:4	--
2	RHA concrete	90%:10%	1:2:4	4.23%
3	RHA concrete	80%:20%	1:2:4	5.93%
4	RHA concrete	70%:30%	1:2:4	14.31%
5	RHA concrete	50%:50%	1:2:4	17.9%
6	RHA concrete	40%:60%	1:2:4	19.74%
7	RHA concrete	30%:70%	1:2:4	23.6%
8	RHA concrete	20%:80%	1:2:4	27.6%

Table-8 Cost comparison between other alternate construction material specimens and conventional Concrete specimen.

SI No	Types of concrete	Prop	%age of saving w.r.t to conventional concrete
1.	Conventional concrete	1:2:4	--
2	Recycled concrete without plasticizer	1:4	28.1%
5.a	Recycled concrete with plasticizer	1:4	25.2%
6.a	Saw dust concrete	1:(1.5+0.5):4	30.1%
7.a	Concrete using brick bat	1:2:4	15.42%

Vehicle's Tracking and Recognition Using a Distributed Surveillance System for Urban Traffic Management

Peyman Babaei

Dep. of Computer, West Tehran Branch, Islamic Azad University, Tehran, Iran

Abstract: This paper proposes an unsupervised vehicle's tracking and recognition methods for urban Traffic surveillance in a distributed cooperative manner. Vehicle's matching in a multi-camera surveillance system is a fundamental issue for increasing the accuracy of recognition. In intelligent transportation systems (ITS), especially in field of urban traffic management, intersections monitoring is one of the critical and challenging tasks. In multi-camera traffic surveillance system, videos have different characteristics such as pose, scale and illumination. Therefore it is necessary to use a hybrid scheme of scale invariant feature transform (SIFT) to detection and recognition vehicle's behavior in multi view more accurately and conveniently. The main focus of this paper is to analyze activities at intersection by a distributed cooperative system for tracking and recognition vehicles to extract traffic flows which assists in regulating traffic lights for using in smart cameras. Extracting the trajectories help to detect abnormal behavior which may be occluded in single-camera surveillance. Distributed cooperative's fundamental purpose is to efficiently reduce the transmission rate and also analyze an intersection scene and report statics and information of interest.

Keywords: Distributed system, Intersection monitoring, Multi-camera surveillance, Vehicle's behaviour learning, urban traffic management.

I. INTRODUCTION

Video surveillance is widely employed in commercial applications and public transportation for purposes of statistics gathering, processing and traffic flow monitoring. The number of cameras and complexity of surveillance systems have been continuously increasing to have better coverage and accuracy. Tracking and behavior recognition are two fundamental tasks in this regard. Multi-camera systems become increasingly attractive in machine vision. Applications include multi view object tracking, event detection, occlusion handling and etc. For many applications, there may be constraints of transmission bandwidth and complexity in analyzing a huge amount of data centrally. In intelligent transportation systems (ITS), the convenient conditions are aroused from autonomous agents making decisions in a decentralized manner. In this paper, we develop method for tracking and recognition by a traffic video surveillance system of two distributed cameras with a partially overlapping field of view. We show how to develop methods for tracking and recognition in a system where processing and decision is distributed across the cameras.

This paper is organized as follows: an overview of the past works in section2. Our proposed architecture and algorithm is presented in section3. Results of subjective

evaluations and objective performance measurements with respect to Ground-truth are presented in section4. Section5 contains the conclusion.

II. PAST WORKS ON MULTI-CAMERA SURVEILLANCE

Features' matching between multiple images of a scene is an important component of many computer vision tasks. In the last few years, a lot of works in detecting, describing and matching feature points has deployed. Although the correspondences can be hand selected, such a procedure is hardly conceivable as the number of cameras increases or when the camera configuration changes frequently, as in a network of pan-tilt-zoom cameras [1]. Other methods for finding correspondences across cameras [2] have been developed through a feature detection method such as the Harris corner detection method [3] or scale invariant feature transform (SIFT) [4]. In [5] shown that corners were efficient for tracking and estimating structure from motion. A corner detector is robust to changes in rotation and intensity but is very sensitive to changes in scale. The Harris detector finds points where the local image geometry has high curvature in the direction of both maximal and minimal curvature, as provided by the eigen-values of the Hessian matrix. They develop an efficient method for determining the relative magnitude of the eigen-values without explicitly computing them. Such color-based matching methods have also been used to track moving objects across cameras [6, 7]. Scale invariant features matching were first proposed in [8] and attracted the attention of the computer vision systems for invariant to scale, rotation, and view-point variations. Also uses a scale-invariant detector in the difference of Gaussian (DOG) scale space. In [4] fits a quadratic to the local scale-space neighborhood to improve accuracy. He then creates a Scale Invariant Feature Transform (SIFT) descriptor to match key-points using a Euclidean distance metric in an efficient best-bin first algorithm where a match is rejected if the ratio of the best and second best matches is greater than a threshold.

A comparative study of many local image descriptors [9] shows the superiority of SIFT with respect to other feature descriptors for the case of several local transformations. In [10] develop a scale-invariant Harris detector that keeps key points at each scale only if it's a maximum in the Laplacian scale-space [11]. More recently, in [12] integrate edge-based features with local feature-based recognition using a structure similar to shape contexts [13] for general object-class recognition. In [14] propose a matching technique based on the Harris corner detector and a description based on the Fourier transform to achieve invariance to rotation. Harris corners are also used in [15], where rotation invariance is obtained by a hierarchal

sampling that starts from the direction of the gradient. In [16] introduce the concept of maximally stable external region to be used for robust matching. These regions are connected components of pixels which are brighter or darker than pixels on the region's contour; they are invariant to affine and perspective transform, and to monotonic transformation of image intensities. Among the many recent works populating the literature on key-point detection, it is worth mentioning the scale and affine invariant interesting points recently proposed in [17], as they appear to be among the most promising key-point detectors to date. The detection algorithm can be sketched as follows: first Harris corners are detected at multiple scales, and then points at which a local measure of variation is maximal over scale are selected. This provides a set of distinctive points at the appropriate scale. Finally, an iterative algorithm modifies location, scale, and neighborhood of each point and converges to affine invariant points. In [18] describe a matching procedure wherein motion trajectories of objects tracked in different cameras are matched so that the overall ground plane can be aligned across cameras following a homograph transformation. A similar approach has been proposed in [19-21] which again motion tracks are matched together. However, although use scene dynamics to find matches, unlike our method, these methods first need to solve the problems of single camera tracking and data association across cameras, which is difficult in highly cluttered scenes or when moving objects occlude each other.

III. PROPOSED ARCHITECTURE AND ALGORITHM

First, we review the function of a typical single-camera and multi-camera surveillance system as presented in our previous work. At that work, as mentioned below the system was centralized. Next, the architecture and algorithm of distributed cooperative system is presented.

A. Single-camera and multi-camera surveillance functionality

As presented in our previous work [22], the function of a typical single-camera surveillance system is illustrated in Fig.1. The first part of the processing flowchart is very general, which is marked "Detecting & Matching Features Extraction Pipeline". This pipeline may produce all target information (pose, scale, illumination, color, shape, etc.), and potentially the description of the scene. The end of the processing pipeline, the vehicle tracking and classification is done.

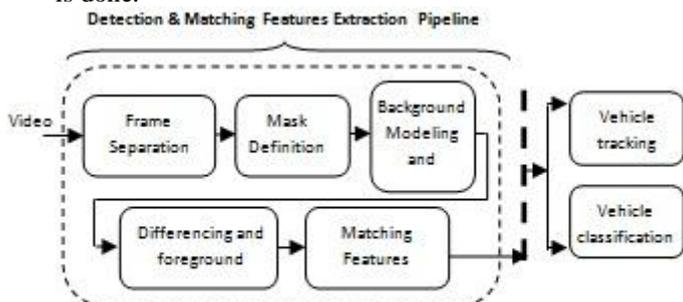


Figure1. Single-camera video surveillance Flowchart

Only the matching features have to be stored, instead of high quality video suitable for automated processing. This method enables the multi-camera surveillance system. The video surveillance system, as described in the above, cannot provide an adequate solution for many applications, Such as urban traffic management with all its associated limitations [23-27]. A multi-camera surveillance system tracking targets from one camera to the next can overcome all these limitations. A typical multi-camera surveillance system is illustrated in Fig.2. Fusing at the matching features level requires merging all the features from the cameras on to a full representation of the environment. This approach distributes the most time consuming processing between the different cameras, and minimizes communication, since only the extracted features needs to be transmitted, no video or image. Given these advantages, system communicates only the matching features for fusion.

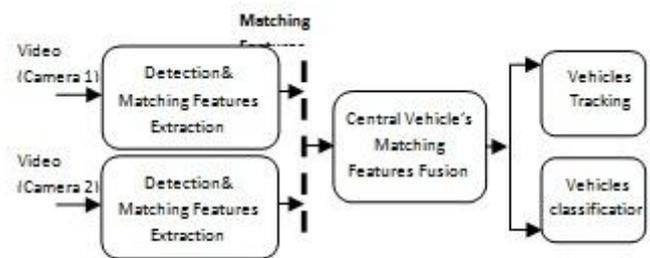


Figure2. Multi-camera video surveillance Flowchart

B. Proposed architecture and algorithm for distributed cooperative system

The problem of multi view activity recognition has been addressed in many papers, but almost the information of multiple views is fused centrally. Our proposed framework is decentralized. The pose of cameras at intersection is shown in Fig.3. In Fig.4, the structure of distributing levels is illustrated.

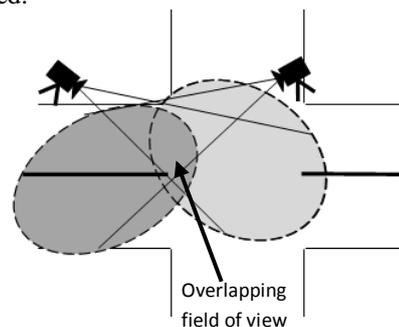


Figure3. Camera setup of cooperative system

Each of the cameras has processing cores in four levels which is described in the flowchart in Fig.4. The input stream is fed to detection level. At the decision level, control commands are issued to classify the detected vehicles based on extracted description features. Processing cores in three upper levels exchange the requisite information to track and recognition more accurately.

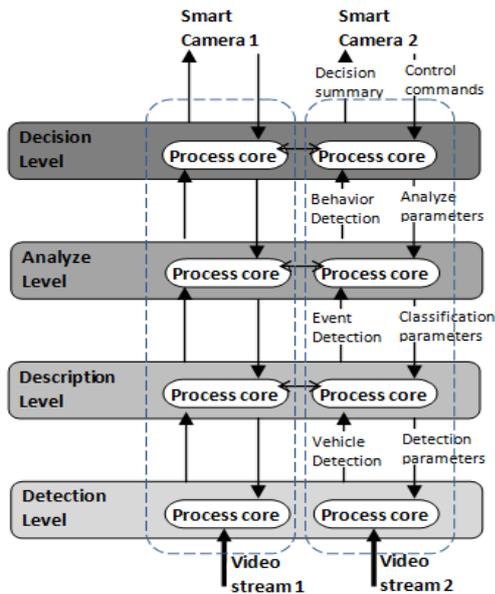


Figure4. Cooperative Levels in proposed distributed system

The principle features of our scheme are summarized in the following:

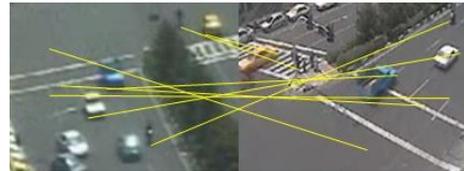
Communication Efficiency: distributed cooperative system is particularly well-suited for low bandwidth; therefore the requirement processing is done locally.

Unsupervised: The method does not require the pre-calibration into the scene and, hence, can be used in traffic scenes where the system administrator may not have control over the activities taking place.

The SIFT (Scale Invariant Feature Transform) [4] has been shown to perform better than other local descriptors [9]. Given a feature point, the SIFT descriptor computes the gradient vector for each pixel in the feature point's neighborhood and builds a normalized histogram of gradient directions. The SIFT descriptor creates a 16x16 neighborhood that is partitioned into 16 sub-regions of 4x4 pixels each. For each pixel within a sub-region, SIFT adds the pixel's gradient vector to a histogram of gradient directions by quantizing each orientation to one of 8 directions and weighting the contribution of each vector by its magnitude. Each gradient direction is further weighted by a Gaussian of scale $\sigma = n/2$ where n is the neighborhood size and the values are distributed to neighboring bins using interpolation to reduce boundary effects as samples move between positions and orientations. Fig.5 shows the matching results using SIFT created for a corresponding pair of points in two intersection scenes.



(a)



(b)

Figure5. Three different intersection scenes, (matching results using SIFT)

IV. EXPERIMENTAL RESULTS

Here it is shown that SIFT lead to excellent performances compared to other existing approaches. As explained, SIFT description is computed as follows: once a key-point is located and its scale has been estimated, one or more orientations are assigned to it based on local image gradient direction around the key-point. Then, image gradient magnitude and orientation are sampled around the key-point, using the scale of the key-point to select the level of Gaussian blur. The gradient orientations obtained are rotated with respect to the key-point orientation previously computed. Finally, the area around the key-point is divided in sub-regions, each of which is associated an orientations histogram weighted with the magnitude. We have experimented with various feature detectors including the Harris corner detector (HCD), curvilinear structure detector (CSD), and difference of Gaussian (DoG) scale space. In Fig.6, the experimental result contain the comparison of these methods is shown.

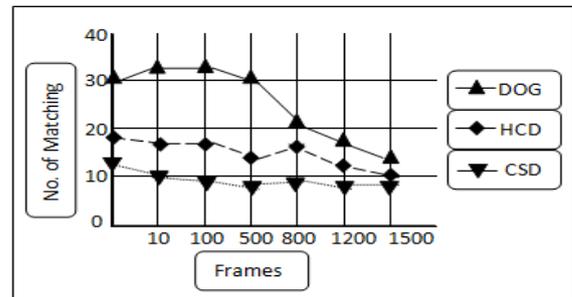


Figure6. Efficiency comparison in intersection traffic scenes

In table1 counting and classification results are presented. As shown, the overall accuracy is about 91% for using DOG detector in counting cars and about 90% for Bus and Trucks. This system can be as an input to calibration system in multi-camera surveillance system.

Table1. Counting and classification results

Vehicle Type	No. of Vehicle's matching in overlapping field of view					
	Bus & Truck			Car		
	DOG	HCD	CSD	DOG	HCD	CSD
Count by Detector	27	21	18	116	108	101
Ground Truth	30	30	30	127	127	127
Precision	90%	70%	60%	91%	85%	79%

V. CONCLUSION

In this paper we considered the problem of features matching in a distributed cooperative system with overlapping fields of view. We showed that using SIFT point descriptors in a distributed cooperative surveillance system can improve the performance with respect to the other calibration systems. In particular it returned good results for scale changes, severe zoom and image plane rotations, and large view-point variations. These conclusions are supported by an extensive experimental evaluation, on different traffic scenes in urban traffic. Therefore, tracking and recognition using SIFT becomes feasible. This should result in highly robust trackers.

ACKNOWLEDGEMENTS

This work was supported by Islamic Azad University, West Tehran Branch.

References

- [1] E.B.Ermis, P.Clarot, P.Jodoin, "Activity Based Matching in Distributed Camera Networks," *IEEE Transaction on Image Processing*, vol. 19, no. 10, OCT. 2010, pp.2595–2613.
- [2] D.Devarajan, Z.Cheng, and R Radke, "Calibrating distributed camera networks," *Proc. IEEE*, vol. 96, no. 10, OCT. 2008, pp. 1625–1639.
- [3] C.Harris and M.Stephens, "A combined corner and edge detector," in *Proc of 4th Alvey Vision Conf.*, 1988, pp. 147–151.
- [4] D.Lowe, "Distinctive image features from scale-invariant keypoints," in *Int. J. Comput. Vis.*, vol. 60, no. 2, 2004, pp. 91–110.
- [5] C. Harris, *Geometry from visual motion*. in: A. Blake, A.Yuille (Eds), *Active Vision*, MIT Press, 1992.
- [6] B.Song and A.R.Chowdhury, "Stochastic adaptive tracking in a camera network," in *Proc.IEEE Int.Conf.Computer Vision*, 2007, pp.1–8.
- [7] S.Khan and M.Shah, "Tracking multiple occluding people by localizing on multiple scene planes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, Mar. 2009, pp. 505–519.
- [8] D.G.Lowe, "Object recognition from local scale-invariant features," in *Proc. Of ICCV*, 1999, pp. 1150–1157.
- [9] K.Mikolajczyk, C.Schmid, "A performance evaluation of local descriptors," in *Proc. Of CVPR*, 2003, pp. 257–263.
- [10] K.Mikolajczyk and C.Schmid, "Indexing based on scale invariant interest points," in *Proc. Of ICCV*, 2001, pp. 525–531.
- [11] T.Lindeberg, "Feature detection with automatic scale selection," in *Proc. Of IJCV*, vol. 30, no.2, 1998, pp.79–116.
- [12] K.Mikolajczyk, A.Zisserman, "Shape recognition with edge-based features," in *Proc. of the British conf. Machine Vision* 2003.
- [13] S.Belongie, J.Malik and J.Puzicha, "Shape context: A new descriptor for shape matching and object recognition," in *Proc. Of NIPS*, 2000 pp. 831–837.
- [14] A.Baumberg, "Reliable feature matching across widely separated views," in *Proc. of CVPR*, 2000, pp.774–781.
- [15] N.Alleazard, M.Dhome, F.Jurie, "Recognition of 3D textured objects by mixing view-based and model based representations," in *Proc. of ICPR*, 2000.
- [16] J.Matas, O.Chum, M.Urban, T.Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *Proc. of BMVC*, 2002, pp. 384–393.
- [17] K.Mikolajczyk, C.Schmid, "Scale and affine invariant interest point detectors," in *Int. J. Com. Vis.* Vol. 60, no.1, 2004, pp. 63–86.
- [18] L.Lee, R.Romano, and G.Stein, "Monitoring activities from multiple video streams: Establishing a common coordinate frame," *IEEE Trans.Pattern Anal.Mach Intell.*, vol.22, no.8, Aug.2000. pp.758–767.
- [19] S.Khan and M.Shah, "Consistent labeling of tracked objects in multiple cameras with overlapping fields of view," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, Oct. 2003, pp.1355–1360.
- [20] X.Wang, K.Tieu, and E.Grimson, "Correspondence-free activity analysis and scene modeling in multiple camera views," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, Jan. 2010, pp. 56–71.
- [21] D.Makris, T.Ellis, and J.Black, "Bridging the gap between cameras," in *Proc. of CVPR*, vol. 2, 2004, pp. 205–210.
- [22] P.Babaei and M.Fathy, "Multi-Camera Systems Evaluation in Urban Traffic Surveillance versus Traditional Single-Camera Systems for Vehicles Tracking," in *Proc. of ICCEE 2010*, vol.4, Nov.2010, pp.438–442.
- [23] H. Veeraraghavan and N. Papanikolopoulos, "Combining multiple tracking modalities for vehicle tracking at traffic intersections," in *IEEE Conf. on Robotics and Automation*, 2004.
- [24] S.Khan and M.Shah, "Consistent labeling of tracking objects in multiple cameras with overlapping fields of view", In *IEEE Trans.*, on PAMI, vol.25, no.10, Oct.2003, pp.1355–1360.
- [25] Q.Zhou, J.Park and J.K.Aggarwal, "Quaternion-Based Tracking of Multiple Objects in synchronized video", In *Proc. of Intl. Sym. on Computer and Information Sciences*, 2003, pp.430–438.
- [26] S.L.Dockstsder and A.M.Tekalp, "Multiple camera tracking of interacting and occluded human motion", In *Proc. of IEEE*. vol.89, no.10, Oct.2001, pp.1441–1455.
- [27] L.Lee, R.Romano and G.Stein, "Monitoring activities from multiple video streams: establishing a common coordinate frame", In *IEEE Trans. on PAMI*, Vol.22, no.8, Aug.2000, pp.758–767.

Applications of Euler's Theorem

To real life and life science

Dr. Mangala Gurjar

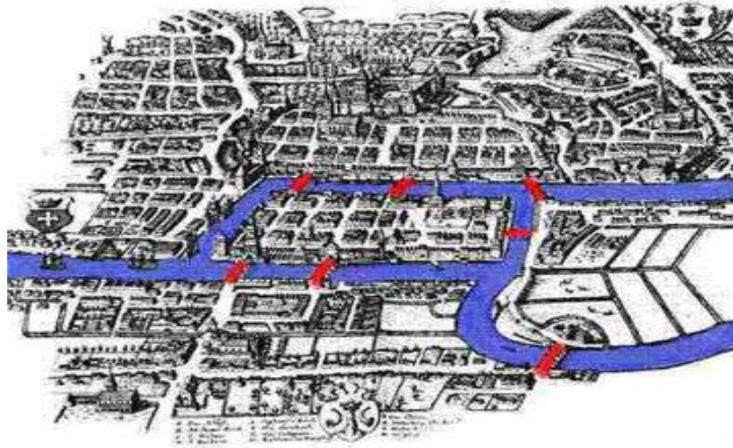
Abstract: In the year 1735 the Swiss mathematician Euler solved the famous seven bridges problem. Euler's solution of the Königsberg's bridges problem is considered to be the first theorem of graph theory which is a branch of combinatorics. This theorem is simple yet has many applications. In the following article we discuss the seven bridges problem followed by some interesting applications of Euler's theorem.

I. Introduction

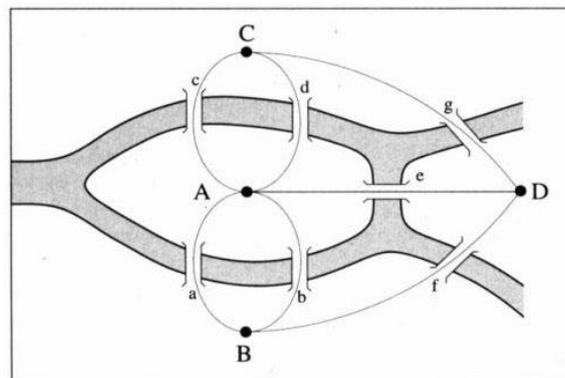
William Wordsworth described Mathematics as an independent world created out of pure intelligence. But even though the creations of Mathematicians seem to have come out of thin air by some people they have their inspirations from some concrete experience. In this paper we will see an example of a concrete problem giving rise to a mathematical theory which again has applications in real life.

Following is the aerial photograph of the town of Königsberg in Russia. There is an island formed due to joining of two rivers. We can also see seven bridges across the rivers. For centuries people wondered whether it was possible to start at a point, walk along all the bridges exactly once and come back to the same point again. They could not do it. But that does not prove that nobody can do it. In the year 1735 Euler used schematic graph to solve the problem. He proved that it was impossible. That was the origin of Graph theory which later found applications in diverse areas. In addition Euler noticed that the key information was the number of bridges and the list of their endpoints (rather than their exact positions).

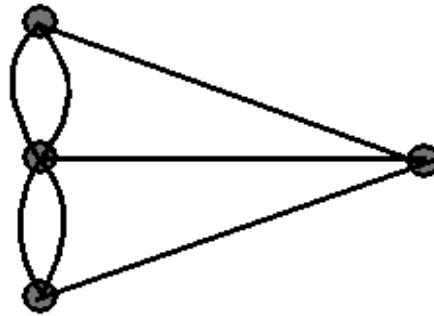
This gave rise to the development of topology:



This can be represented as:



Which can be further simplified to:



Here the vertices are the land masses shrunk to a point and lines or edges are the bridges.

Because no edge is allowed to be repeated every vertex must be entered and exited through different edges. In order to do that the degree of each vertex which is the number of lines originating at the vertex must be even. But we see that degree of each vertex is three or five which is odd. Hence it is impossible to go along all the edges exactly once and complete the circuit.

This simple observation made by Euler that a circuit can be completed including all the edges exactly once if and only if the degree of each vertex is even is called as Euler's theorem. It can be applied to many different situations. To apply Euler's theorem we have to only identify the vertices and edges. In different problems different entities become vertices and edges.

For example: Suppose you have a pack of dominos each consisting of two squares and each square containing different number of dots from the set $\{1, 2, 3, 4, \dots, n\}$. In all there are $n(n+1)/2$ different dominos. Can they be arranged in a circle in such a way that squares containing same number of dots are adjacent to each other?

In order to apply graph theory to solve this problem we consider dominos as edges and numbers 1, 2, 3, etc. as vertices. All the dominos will form a complete graph where every vertex is joined to every other vertex. Thus degree of each vertex is $n-1$. Using Euler's theorem a circuit can be completed including all the edges exactly once if and only if $n-1$ is even which means n is odd. In conclusion they can be arranged in a circle if and only if n is odd.

We now turn our attention to a 20th century application of Eulerian Graphs in the subject of DNA-recovery. DNA is a chain of four possible chemicals popularly called as A, C, T, G. The enzymes that break the chain after each G link are called G- fragments and those that break the chain after each G or T link are called T, C-fragments. The problem is that of recovering the original DNA- chain given the sets of its G- fragments and T,C- fragments.

For example: The chain ACCGGATCGTTCTGTG has

G-fragments = {ACCG, G, ATCG, TTCG, TG} and
 T,C- fragments = {AC, C, GGAT, C, GT, T, C, GT, G}

Without using Euler's Theorem if we try the permutations then we have a large number of possibilities for the original sequence. But using Euler's theorem we can reduce the number of possibilities significantly as explained below.

First subject the G- fragments to T, C- fragments and subject the T, C-fragments to G- fragments thus getting the following two sets of sub fragments.

{AC.C.G, G, AT.C.G, T.T.C.G, T.G} and {AC, C, G.G. AT, C, G.T, T, C, G.T, G}

From these we collect interior sub fragments in to set-1 and fragments having only one sub fragments in to set-2. In this case,

Set-1 = {C,C, T, C, G} and Set-2 = {G, AC, C, C, T, C, G}

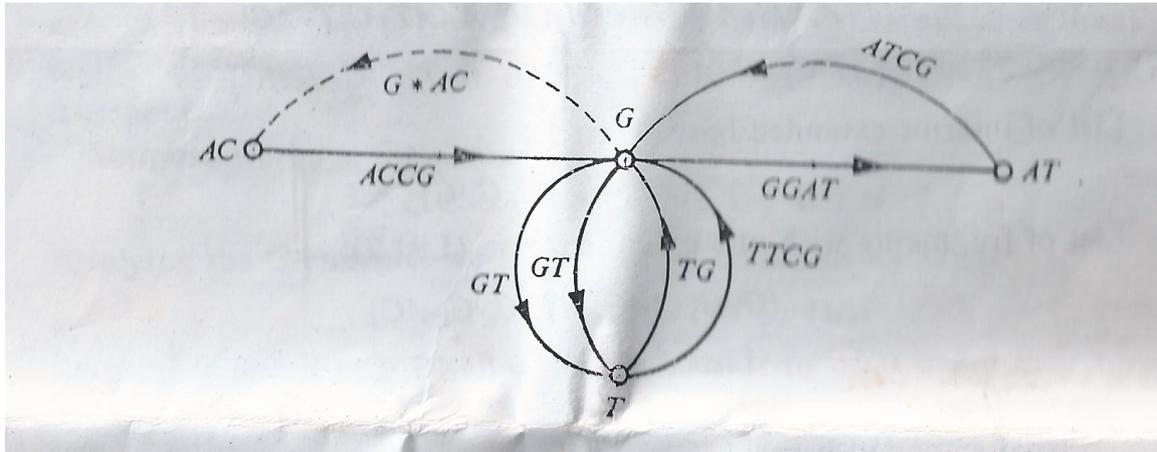
Set-2 will always be super set of Set-1. and Set-2 – Set-1 gives the first and the last sub fragment of the original DNA sequence. In this case Set-2 – Set-1 = {G, AC}.

Out of the two elements of this set the abnormal fragment will be the last.

Abnormal fragment is a G- fragment not ending in G or a T, C fragment not ending in T or C. In this case G is abnormal because it is a T, C fragment not ending in T or C. Thus the original chain must start with AC and end with G.

Now we build a graph with all the fragments having more than one sub fragment as the edges. In this case the set is

{ACCG, ATCG, TTCG, TG, GGAT, GT,GT}. The set of vertices will be identified as we build the graph.



In this graph there are 12 paths starting at AC and ending at G containing all the edges only once. We can add the last edge from G to AC to complete the circuit. Each of these circuits generates a sequence. One among them is the original sequence. Note that the graph constructed above has Eulerian path because the degree of each vertex is even.

Conclusion

The use of Graph Theory has reduced the number of cases to 12 from hundreds of possible cases in DNA recovery.

References

1. Basic Graph Theory By K.R. Parthasarathy –Tata McGraw-Hill Publishing Company Limited(1994).
2. Art and Craft of problem solving By Paul Zeitz John Wiley and sons INC (1999).
3. Culture, Excitement and Relevance of Mathematics By V. Krishnamurthy Wiley Eastern Limited(1990).

Address Dr. Mangala R. Gurjar, M.Sc, Ph.D.(Mathematics)
 Associate Professor & HoD, Mathematics department,
 St. Xavier's College, Mumbai-400001.

Composite entire and meromorphic functions and their growth analysis in the light of order and weak type

Sanjib Kumar Datta¹, Tanmay Biswas², Manab Biswas³

¹(Dept. of Mathematics, University of Kalyani, Kalyani, Dist-Nadia, West Bengal, India)

²(Rajbari, Rabindrapalli, R. N. Tagore Road, P.O. Krishnagar, Dist.- Nadia, West Bengal, India)

³(Barabilla High School, P.O. Haptiagach, Dist-Uttar Dinajpur, West Bengal, India)

Abstract: In this paper we study the growth properties of composite entire and meromorphic functions which improve some earlier results.

AMS Subject Classification (2010): 30D30, 30D35.

Keywords and phrases: Meromorphic function, entire function, composition, order (lower order), growth.

I. Introduction, Definitions and Notations

We denote by \mathbb{C} the set of all finite complex numbers. Let f be a meromorphic function and g be an entire function defined on \mathbb{C} . We use the standard notations and definitions in the theory of entire and meromorphic functions which are available in [4] and [10]. In the sequel we use the following notations:

$\log^{[k]} x = \log (\log^{[k-1]} x)$ for $k = 1, 2, 3, \dots$ and

$\log^{[0]} x = x$;

and

$\exp^{[k]} x = \exp (\exp^{[k-1]} x)$ for $k = 1, 2, 3, \dots$ and

$\exp^{[0]} x = x$.

Definition 1 The order ρ_f and lower order λ_f of an entire function f are defined as

$$\rho_f = \limsup_{r \rightarrow \infty} \frac{\log^{[2]} M(r, f)}{\log r} \text{ and } \lambda_f = \liminf_{r \rightarrow \infty} \frac{\log^{[2]} M(r, f)}{\log r} .$$

If f is meromorphic then

$$\rho_f = \limsup_{r \rightarrow \infty} \frac{\log T(r, f)}{\log r} \text{ and } \lambda_f = \liminf_{r \rightarrow \infty} \frac{\log T(r, f)}{\log r} .$$

The following definition is also well known :

Definition 2 [3] The weak type τ_f of a meromorphic function f of finite positive lower order λ_f is defined by

$$\tau_f = \liminf_{r \rightarrow \infty} \frac{T(r, f)}{r^{\lambda_f}} .$$

For entire f ,

$$\tau_f = \liminf_{r \rightarrow \infty} \frac{\log M(r, f)}{r^{\lambda_f}}, \quad 0 < \lambda_f < \infty .$$

Similarly one can define the growth indicator $\bar{\tau}_f$ of a meromorphic function f of finite positive lower order λ_f as

$$\bar{\tau}_f = \limsup_{r \rightarrow \infty} \frac{T(r, f)}{r^{\lambda_f}} .$$

When f is entire, it can be easily verified that

$$\bar{\tau}_f = \limsup_{r \rightarrow \infty} \frac{\log M(r, f)}{r^{\lambda_f}}, \quad 0 < \lambda_f < \infty .$$

Definition 3 [9] A function $\rho_f(r)$ is called a proximate order of f relative to $T(r, f)$ if

(i) $\rho_f(r)$ is non-negative and continuous for $r \geq r_0$, say,

(ii) $\rho_f(r)$ is differentiable for $r \geq r_0$ except possibly at isolated points at which $\rho_f'(r-0)$ and $\rho_f'(r+0)$ exist,

(iii) $\lim_{r \rightarrow \infty} \rho_f(r) = \rho_f < \infty$,

(iv) $\lim_{r \rightarrow \infty} r \rho_f'(r) \log r = 0$ and

(v) $\limsup_{r \rightarrow \infty} \frac{T(r, f)}{r^{\rho_f(r)}} = 1$.

In the line of Definition 3 the following definition may be given :

Definition 4 A function $\lambda_f(r)$ is called a lower proximate order of f relative to $T(r, f)$ if

(i) $\lambda_f(r)$ is non-negative and continuous for $r \geq r_0$, say,

(ii) $\lambda_f(r)$ is differentiable for $r \geq r_0$ except possibly at isolated points at which $\lambda_f'(r-0)$ and $\lambda_f'(r+0)$ exist,

(iii) $\lim_{r \rightarrow \infty} \lambda_f(r) = \lambda_f < \infty$,

(iv) $\lim_{r \rightarrow \infty} r \lambda_f'(r) \log r = 0$ and

(v) $\liminf_{r \rightarrow \infty} \frac{T(r, f)}{r^{\lambda_f(r)}} = 1$.

In the paper we establish some newly developed results based on the comparative growth properties of composite entire or meromorphic functions.

II. Lemmas.

In this section we present some lemmas which will be needed in the sequel.

Lemma 1 [1] Let f be meromorphic and g be entire. Then for all sufficiently large values of r ,

$$T(r, f \circ g) \leq \{1 + o(1)\} \frac{T(r, g)}{\log M(r, g)} T(M(r, g), f).$$

Lemma 2 [2] Let f be meromorphic and g be entire and suppose that $0 < \mu < \rho_g \leq \infty$. Then for a sequence of values of r tending to infinity,

$$T(r, f \circ g) \geq T(\exp(r^\mu), f).$$

Lemma 3 [6] Let f be meromorphic and g be entire such that $0 < \mu < \rho_g \leq \infty$ and $\lambda_f > 0$. Then for a sequence of values of r tending to infinity,

$$T(r, f \circ g) \geq T(\exp(r^\mu), g).$$

Lemma 4 [5] If f be an entire function then for $\delta (> 0)$ the function $r^{\rho_f + \delta - \rho_f(r)}$ is ultimately an increasing function of r .

Lemma 5 [7] Let f be an entire function. Then for $\delta (> 0)$ the function $r^{\lambda_f + \delta - \lambda_f(r)}$ is ultimately an increasing function of r .

III. Theorems.

In this section we present the main results of the paper.

Theorem 1 Let f, h be any two meromorphic functions and g, k be any two entire functions such that $\rho_h < \infty, \rho_k < \rho_g$ and $\lambda_f > 0$. Then

$$\liminf_{r \rightarrow \infty} \frac{\log \{T(r, h \circ k) \log M(r, k)\}}{\log T(r, f \circ g)} = 0.$$

Proof. As $\rho_g < \rho_k$, we can choose $\varepsilon (> 0)$ in such a way that

$$\rho_k + \varepsilon < \rho_g - \varepsilon < \rho_g. \quad (1)$$

Now from (1) and Lemma 2 it follows that for a sequence of values of r tending to infinity that

$$\log T(r, fog) \geq \log T(\exp r^{(\rho_g - \varepsilon)}, h)$$

$$i. e., \log T(r, fog) \geq (\lambda_f - \varepsilon) \log \exp r^{(\rho_g - \varepsilon)}$$

$$i. e., \log T(r, fog) \geq (\lambda_f - \varepsilon) r^{(\rho_g - \varepsilon)}. \quad (2)$$

Again we have from Lemma 1 for all sufficiently large values of r ,

$$T(r, hok) \log M(r, k) \leq \{1 + o(1)\} T(r, k) T(M(r, k), h)$$

$$i. e., \log \{T(r, hok) \log M(r, k)\}$$

$$\leq (\rho_k + \varepsilon) \log r + (\rho_h + \varepsilon) \log M(r, k) + O(1)$$

$$i. e., \log \{T(r, hok) \log M(r, k)\}$$

$$\leq (\rho_k + \varepsilon) \log r + (\rho_h + \varepsilon) r^{(\rho_k + \varepsilon)} + O(1). \quad (3)$$

Therefore from (2) and (3) we obtain for a sequence of values of r tending to infinity that

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} \leq \frac{(\rho_k + \varepsilon) \log r + (\rho_h + \varepsilon) r^{(\rho_k + \varepsilon)} + O(1)}{(\lambda_f - \varepsilon) r^{(\rho_g - \varepsilon)}}. \quad (4)$$

Now in view of (1) it follows from (4) that

$$\liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = 0.$$

This proves the theorem.

Remark 1 For the validity of Theorem 1, the conditions $\rho_h < \infty$, $\rho_k < \rho_g$ and $\lambda_f > 0$ are necessary but for meromorphic h with order zero Theorem 1 also holds for $\rho_g \leq \rho_k$ which are evident from the following examples:

Example 1 Let $f = k = \exp z$, $g = \exp(z^2)$ and $h = \exp^{[2]} z$.

Then $\lambda_f = 1 > 0$, $\rho_h = \infty$ and $\lambda_k = \rho_k = 1 < 2 = \rho_g$.

Now

$$T(r, fog) \leq \log M(r, fog) = \exp(r^2)$$

$$\text{and } 3T(2r, hok) \geq \log M(r, hok) = \exp^{[2]} r.$$

So

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \frac{\log T(r, hok) + \log^{[2]} M(r, k)}{\log T(r, fog)}$$

$$\geq \frac{\exp \frac{r}{2} + \log r + O(1)}{r^2}$$

$$i. e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \infty.$$

Example 2 Suppose $f = h = g = \exp z$ and $k = \exp(z^2)$.

Then $\rho_f = \lambda_h = \rho_h = \lambda_g = \rho_g = 1$ and $\rho_k = 2$.

Now

$$3T(2r, hok) \geq \log M(r, hok) = \exp(r^2) = r^2$$

$i. e., \log T(r, hok) \geq \frac{r^2}{4} + O(1) .$

Also

$$T(r, fog) \sim \frac{\exp r}{(2\pi^3 r)^{\frac{1}{2}}} .$$

Therefore

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \frac{\log T(r, hok) + \log^{[2]} M(r, k)}{\log T(r, fog)} \geq \frac{\frac{r^2}{4} + O(1) + 2 \log r}{r - \frac{1}{2} \log r + O(1)} .$$

$i. e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \infty .$

Example 3 Suppose $f = z, g = \exp(z^2)$ and $h = k = \exp z .$

Then $\lambda_f = \rho_f = 0 < \infty, \lambda_h = \rho_h = \lambda_k = \rho_k = 1 < 2 = \rho_g .$

Therefore

$$T(r, fog) \leq \log M(r, fog) = r^2$$

$i. e., \log T(r, fog) \leq 2 \log r .$

Also

$$T(r, hok) \sim \frac{\exp r}{(2\pi^3 r)^{\frac{1}{2}}} .$$

Thus

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \frac{\log T(r, hok) + \log^{[2]} M(r, k)}{\log T(r, fog)} \geq \frac{r - \frac{1}{2} \log r + \log r + O(1)}{2 \log r}$$

$i. e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \infty .$

Example 4 Let $f = g = \exp z, h = z$ and $k = \exp(z^2) .$

Then $\rho_f = \rho_g = 1, \lambda_h = \rho_h = 0$ and $\lambda_k = \rho_k = 2 .$

Now

$$T(r, hok) \leq \log M(r, hok) = \log \exp(r^2) = r^2$$

and $T(r, fog) \sim \frac{\exp r}{(2\pi^3 r)^{\frac{1}{2}}} .$

So

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \frac{\log T(r, hok) + \log^{[2]} M(r, k)}{\log T(r, fog)} \leq \frac{4 \log r}{r - \frac{1}{2} \log r + O(1)}$$

$i. e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = 0 .$

Example 5 Let $f = g = k = \exp z$ and $h = z .$

Then $\rho_f = \rho_g = 1, \lambda_h = \rho_h = 0$ and $\lambda_k = \rho_k = 2 .$

Now

$$T(r, hok) \leq \log M(r, hok) = r$$

and $T(r, fog) \sim \frac{\exp r}{(2\pi^3 r)^{\frac{1}{2}}} .$

So

$$\frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = \frac{\log T(r, hok) + \log^{[2]} M(r, k)}{\log T(r, fog)} \leq \frac{2 \log r}{r - \frac{1}{2} \log r + O(1)}$$

$$i. e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, hok) \log M(r, k)\}}{\log T(r, fog)} = 0.$$

In the line of Theorem 1 one can easily prove the following theorem :

Theorem 2 Let f, h be any two meromorphic functions and g, k be any two entire functions with $\rho_h < \infty, \rho_k < \rho_g$ and $\lambda_f > 0$. Then

$$\liminf_{r \rightarrow \infty} \frac{\log^{[2]} \{T(r, hok) \log M(r, k)\}}{\log^{[2]} T(r, fog)} \leq \frac{\rho_k}{\rho_g}.$$

The proof is omitted .

In the line of Theorem 2 the following corollary may be deduced :

Corollary 1 Let f, h be meromorphic and g, k be entire such that $\rho_h < \infty, \rho_k < \rho_g$ and $\lambda_f > 0$. Then

$$\liminf_{r \rightarrow \infty} \frac{\log^{[3]} \{T(r, hok) \log M(r, k)\}}{\log^{[3]} T(r, fog)} \leq 1.$$

Theorem 3 Let f, h be meromorphic and g, k be entire such that (i) $\rho_f < \infty$, (ii) $\lambda_h > 0$, (iii) $\lambda_k > 0$, (iv) $\lambda_g < \rho_k$ and (v) $0 < \lambda_g < \infty, \bar{\tau}_g < \infty$. Then

$$\liminf_{r \rightarrow \infty} \frac{\log \{T(r, fog) \log M(r, g)\}}{\log T(r, hok)} \leq \rho_f \bar{\tau}_g \cdot \min \{ \lambda_h^{-1}, \lambda_k^{-1} \}.$$

Proof. By Lemma 1 we obtain for all sufficiently large values of r ,

$$T(r, fog) \log M(r, g) \leq \{1 + o(1)\} T(r, g) T(M(r, g), f)$$

$$i. e., \log \{T(r, fog) \log M(r, g)\}$$

$$\leq (\rho_g + \varepsilon) \log r + (\rho_f + \varepsilon) \log M(r, g) + O(1)$$

$$i. e., \log \{T(r, fog) \log M(r, g)\}$$

$$\leq (\rho_g + \varepsilon) \log r + (\rho_f + \varepsilon) (\bar{\tau}_g + \varepsilon) r^{\lambda_g} + O(1). (5)$$

Since $\lambda_g < \rho_k$, in view of Lemma 2 it follows for a sequence of values of r tending to infinity that

$$\log T(r, hok) \geq \log T(\exp(r^{\lambda_g}), h)$$

$$i. e., \log T(r, hok) \geq (\lambda_h - \varepsilon) \log \exp(r^{\lambda_g})$$

$$i. e., \log T(r, hok) \geq (\lambda_h - \varepsilon) r^{\lambda_g}. \quad (6)$$

Similarly in view of Lemma 3 we have for a sequence of values of r tending to infinity

$$\log T(r, hok) \geq \log T(\exp(r^{\lambda_g}), k)$$

$$i. e., \log T(r, hok) \geq (\lambda_k - \varepsilon) \log \exp(r^{\lambda_g})$$

$$i. e., \log T(r, hok) \geq (\lambda_k - \varepsilon) r^{\lambda_g}, \quad (7)$$

where $0 < \varepsilon < \min \{ \lambda_h, \lambda_k \}$.

Now from (5) and (6) we obtain for a sequence of values of r tending to infinity that

$$\frac{\log \{T(r, fog) \log M(r, g)\}}{\log T(r, hok)} \leq \frac{(\rho_g + \varepsilon) \log r + (\rho_f + \varepsilon) (\bar{\tau}_g + \varepsilon) r^{\lambda_g} + O(1)}{(\lambda_h - \varepsilon) r^{\lambda_g}}$$

$$i.e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, fog) \log M(r, g)\}}{\log T(r, hok)} \leq \frac{\rho_f \bar{\tau}_g}{\lambda_h} . (8)$$

Analogously from (5) and (7) it follows for a sequence of values of r tending to infinity that

$$\frac{\log \{T(r, fog) \log M(r, g)\}}{\log T(r, hok)} \leq \frac{(\rho_g + \varepsilon) \log r + (\rho_f + \varepsilon)(\bar{\tau}_g + \varepsilon)r^{\lambda_g} + O(1)}{(\lambda_k - \varepsilon)r^{\lambda_g}}$$

$$i.e., \liminf_{r \rightarrow \infty} \frac{\log \{T(r, fog) \log M(r, g)\}}{\log T(r, hok)} \leq \frac{\rho_f \bar{\tau}_g}{\lambda_k} . (9)$$

Thus the theorem follows from (8) and (9) .

In the line of Theorem 3 one can easily prove the following theorem :

Theorem 4 Let f, h be meromorphic and g, k be entire such that (i) $\rho_h < \infty$, (ii) $\lambda_f > 0$, (iii) $\lambda_g > 0$, (iv) $\lambda_k > \rho_g$ and (v) $0 < \lambda_k < \infty, \bar{\tau}_k < \infty$. Then

$$\limsup_{r \rightarrow \infty} \frac{\log T(r, fog)}{\log \{T(r, hok) \log M(r, k)\}} \geq (\rho_h \bar{\tau}_k)^{-1} \cdot \max \{ \lambda_f, \lambda_g \} .$$

The proof is omitted .

Theorem 5 Let f be a meromorphic function and g, h be two entire functions such that $\rho_g < \infty, \rho_f < \infty$ and $\lambda_h > 0$. Then for any $\alpha > 1$

$$\liminf_{r \rightarrow \infty} \frac{\log T(r, fog)}{\log T(r, hog)} \leq \left(\frac{\alpha+1}{\alpha-1}\right) \cdot \frac{\rho_f}{\lambda_h} \cdot (4\alpha)^{\rho_g} .$$

Proof. Since $T(r, g) \leq \log^+ M(r, g)$, we obtain by Lemma 1 for $\varepsilon (> 0)$ and for all sufficiently large values of r ,

$$T(r, fog) \leq \{1 + o(1)\} T(M(r, g), f)$$

$$i.e., \log T(r, fog) \leq (\rho_f + \varepsilon) \log M(r, g) + O(1) . (10)$$

For all sufficiently large values of r we know that

$$T(r, hog) \geq \frac{1}{3} \log M \left\{ \frac{r}{4}, g \right\} + o(1), h \} \{ cf. [8] \}$$

For $\varepsilon (0 < \varepsilon < \min \{ \lambda_h, \lambda_k \})$ we get for all sufficiently large values of r ,

$$\log T(r, hog) \geq (\lambda_h - \varepsilon) \log \left\{ \frac{1}{8} M \left(\frac{r}{4}, g \right) + o(1) \right\} + O(1)$$

$$i.e., \log T(r, hog) \geq (\lambda_h - \varepsilon) \log \left\{ \frac{1}{9} M \left(\frac{r}{4}, g \right) \right\} + O(1)$$

$$i.e., \log T(r, hog) \geq (\lambda_h - \varepsilon) \log M \left(\frac{r}{4}, g \right) + O(1)$$

$$i.e., \log T(r, hog) \geq (\lambda_h - \varepsilon) T \left(\frac{r}{4}, g \right) + O(1) . (11)$$

Since $\varepsilon (> 0)$ is arbitrary, it follows from (10) and (11) for all sufficiently large values of r ,

$$\liminf_{r \rightarrow \infty} \frac{\log T(r, fog)}{\log T(r, hog)} \leq \frac{\rho_f}{\lambda_h} \liminf_{r \rightarrow \infty} \frac{\log M(r, g)}{T \left(\frac{r}{4}, g \right)} (12)$$

Since $\limsup_{r \rightarrow \infty} \frac{T(r, g)}{r^{\rho_g(r)}} = 1$, for given $\varepsilon (0 < \varepsilon < 1)$ we get for all sufficiently large values of r ,

$$T(r, g) < (1 + \varepsilon) r^{\rho_g(r)} (13)$$

and for a sequence of values of r tending to infinity

$$T(r, g) > (1 - \varepsilon) r^{\rho_g(r)} . (14)$$

Since for any $\alpha > 1$, $\log M(r, g) \leq \frac{\alpha+1}{\alpha-1} T(\alpha r, g)$, in view of (13), (14) and for any $\delta (> 0)$ we get for a sequence of values of r tending to infinity that

$$\begin{aligned} \frac{\log M(r, g)}{T\left(\frac{r}{4}, g\right)} &\leq \frac{\frac{\alpha+1}{\alpha-1} (1 + \varepsilon)}{(1 - \varepsilon)} \cdot \frac{(\alpha r)^{\rho_g + \delta}}{(\alpha r)^{\rho_g + \delta - \rho_g(\alpha r)}} \cdot \frac{1}{\left(\frac{r}{4}\right)^{\rho_g \left(\frac{r}{4}\right)}} \\ &\leq \left(\frac{\alpha + 1}{\alpha - 1}\right) \frac{(1 + \varepsilon)}{(1 - \varepsilon)} \cdot \frac{\left(\frac{4\alpha r}{4}\right)^{\rho_g + \delta}}{\left(\frac{4\alpha r}{4}\right)^{\rho_g + \delta - \rho_g \left(\frac{4\alpha r}{4}\right)}} \cdot \frac{1}{\left(\frac{r}{4}\right)^{\rho_g \left(\frac{r}{4}\right)}} \\ &\leq \left(\frac{\alpha + 1}{\alpha - 1}\right) \frac{(1 + \varepsilon)}{(1 - \varepsilon)} \cdot (4\alpha)^{\rho_g + \delta} \end{aligned}$$

because $r^{\rho_g + \delta - \rho_g(r)}$ is ultimately an increasing function of r . Since $\varepsilon (> 0)$ and $\delta (> 0)$ are arbitrary, we obtain that

$$\liminf_{r \rightarrow \infty} \frac{\log M(r, g)}{T\left(\frac{r}{4}, g\right)} \leq \left(\frac{\alpha+1}{\alpha-1}\right) \cdot (4\alpha)^{\rho_g} . \quad (15)$$

Thus from (12) and (15) it follows that

$$\liminf_{r \rightarrow \infty} \frac{\log T(r, f \circ g)}{\log T(r, h \circ g)} \leq \left(\frac{\alpha+1}{\alpha-1}\right) \cdot \frac{\rho_f}{\lambda_h} \cdot (4\alpha)^{\rho_g} .$$

In the line of Theorem 5 one can easily prove the following theorem using the definition of lower proximate order :

Theorem 6 Let f be a meromorphic function and g, h, k be any three entire functions such that $\rho_g < \lambda_k < \infty$ and $\lambda_h < \infty$. Then for any $\alpha > 1$

$$\liminf_{r \rightarrow \infty} \frac{\log T(r, f \circ g)}{\log T(r, h \circ g)} \leq \left(\frac{\alpha+1}{\alpha-1}\right) \cdot \frac{\rho_f}{\lambda_h} \cdot (4\alpha)^{\lambda_g} .$$

The proof is omitted .

References

- [1] Bergweiler, W : On the Nevanlinna Characteristic of a composite function , Complex variables , Vol. 10(1988) , pp. 225-236 .
- [2] Bergweiler, W : On the growth rate of composite meromorphic functions , Complex Variables, Vol. 14 (1990) , pp. 187-196 .
- [3] Datta, S.K. and Jha, A. : On the weak type of meromorphic functions , Int. Math. Forum. , Vol. 4, No. 12 (2009) , pp.569-579 .
- [4] Hayman, W.K : Meromorphic Functions , The Clarendon Press , Oxford , 1964 .
- [5] Lahari, I. : Growth of composite integral functions , Indian J. Pure Appl. Math., Vol. 20, No. 9(1989) , pp. 899-907 .
- [6] Lahiri, I. and Sharma, D.K. : Growth of composite entire and meromorphic functions, Indian J. Pure Appl. Math., Vol. 26 , No. 5(1995) , pp. 451-458 .
- [7] Lahiri, I. and Datta, S.K. : On the growth properties of composite entire and meromorphic functions , Bull. Allahabad Math. Soc., Vol. 18 (2003), pp.15-34 .
- [8] Niino, K. and Yang, C.C. : Some growth relationships on factors of two composite entire functions, Factorization Theory of Meromorphic Functions and Related topics, Marcel Dekker, Inc.(New York and Basel), (1982), pp. 95-99.
- [9] Shah, S.M. : On proximate order of integral functions , Bull. Amer. Math. Soc., Vol. 52(1946), pp. 326-328 .
- [10] Valiron, G. : Lectures on the General Theory of Integral Functions , Chelsea Publishing Company , (1949) .

Finite Element Modeling for Replacement of C.I. Pulley with Suitable Material

Deepak Singathia, Dr. M. L. Aggarwal

*PG Student, Department of Mechanical Engineering, YMCA University of Science and Technology, India
Professor & Dean Academic Affairs, Department of Mechanical Engineering, YMCA University of Science and Technology, India*

Abstract: FEA is used in many analysis problems like structural analysis, thermal analysis, computational fluids dynamics problems etc. In this paper the maximum stress developed in a flat pulley is determined for different materials using ANSYS software. The maximum stress developed i.e. Von Mises stress is then compared for four materials in order to select best material for pulleys. The simulation stress results were compared and factor of safety was calculated on the basis of stress developed. In this present work, the stress analysis of Alumina, Al alloy 1460, Polycarbonate molded pulley is carried out to replace presently used C.I. pulley.

Keywords: Finite Element Analysis, Pulley, Stress.

I. INTRODUCTION

It is necessary for the designer to know the stress distribution in order to prevent failure. In the present work the developed stress in the pulley is determined using FEA with the help of ANSYS. Stresses near the hub end of pulley were evaluated by varying the materials [1]. The finite element analysis was carried out using ANSYS. The results of the finite element analysis were verified experimentally. There are many causes of pulley failure. Among them the maximum bending stress induced near the hub end is one of the causes of pulley failure. Belt-drive mode is presented by simulating the dynamic response of the complex, serpentine belt-drive with tensioners, which are common in automotive engines [2]. A lot of work has been done on pulleys but still C.I. pulleys are still used for heavy load applications. The major object of this paper is to determine: (i) Von Mises Stress developed in the pulley model for the different materials, (ii) To compare the results obtained from ANSYS so as to select best replacement of presently used C.I. pulley for heavy load applications.

II. GEOMETRIC DIMENSION OF PULLEY

The same pulley model is used for all the materials that are used in this analysis. The dimensions of pulley were as follows:

Diameter of pulley	D = 55 mm
Hub Diameter	d = 17 mm
Thickness	T = 9.5 mm
Hub Projection	H = 6.4 mm
Bore Size	b = 5 mm

III. STRESS ANALYSIS OF PULLEY

Total number of element generated in pulley by ANSYS software is 29237 approximately and the element type considered is solid 45. Boundary conditions are applied on the inner side of pulley, i.e. it is fixed and load is applied on periphery of rim. Fig. 1-4 shows the procedure of stress analysis. Fig. 5-6 shows the Von Mises stress and strain developed for the polycarbonate molded. Fig. 7-8 shows the Von Mises stress and strain developed for the Gray Cast Iron. Fig. 9-10 shows the Von-Mises stress and strain developed for the Alumina. Fig. 11-12 shows the Von Mises stress and strain developed for the Al alloy,1460 pressed profile.

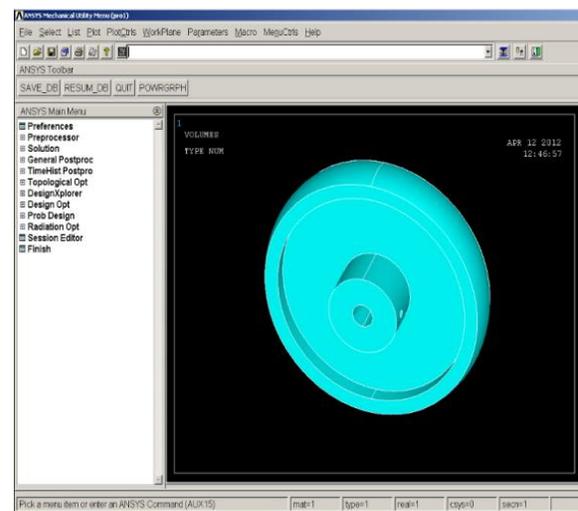


Fig. 1: Imported view of pulley in ANSYS

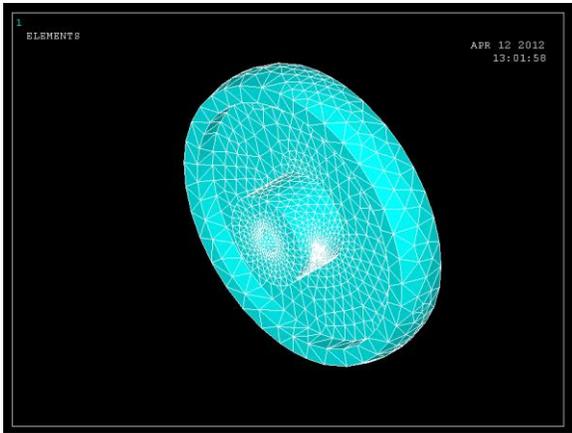


Fig. 2: Meshing

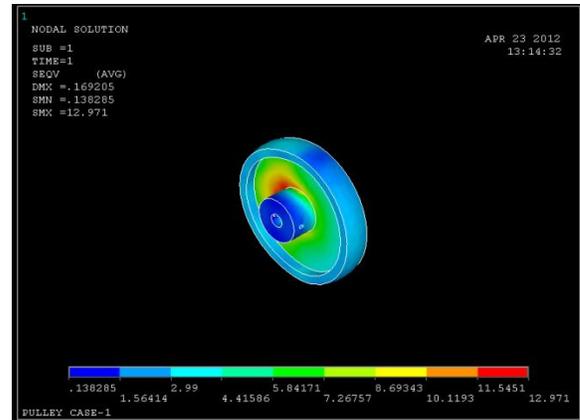


Fig. 5: Von Mises stress developed in Polycarbonate

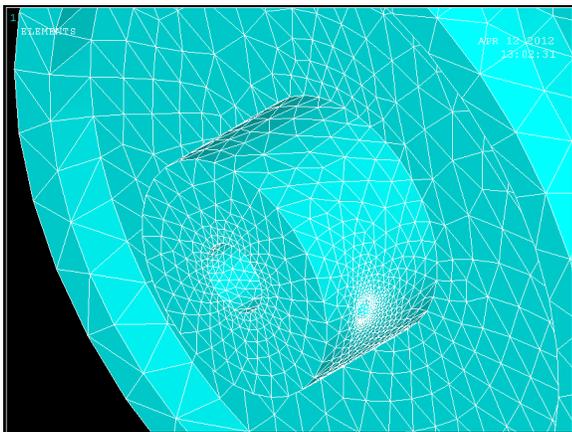


Fig. 3: Meshing closer view

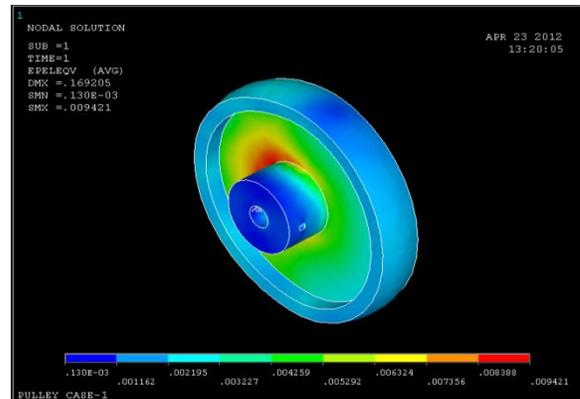


Fig. 6: Von Mises strain developed in Polycarbonate

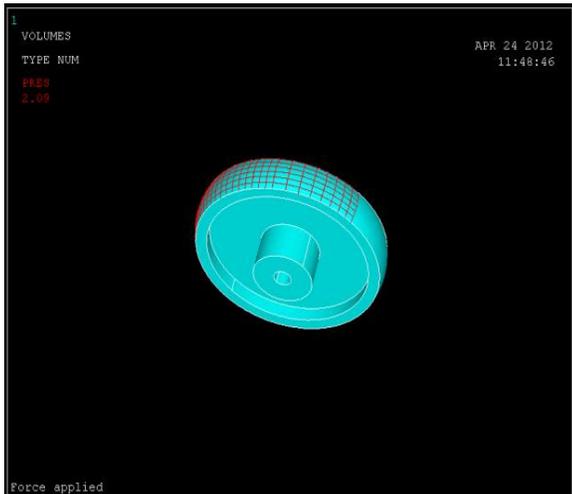


Fig. 4: Load case applied

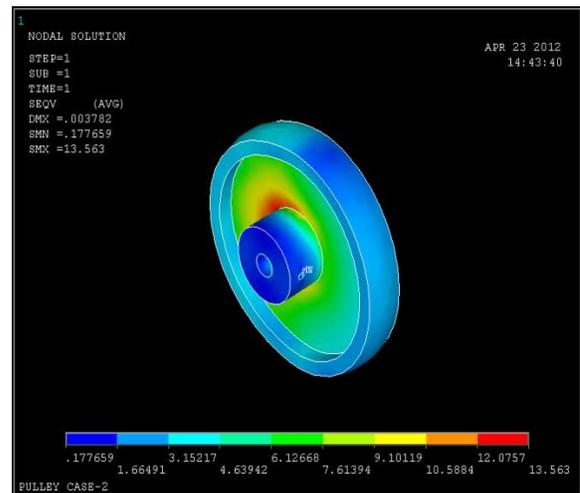


Fig. 7: Von Mises stress developed in Gray Cast Iron

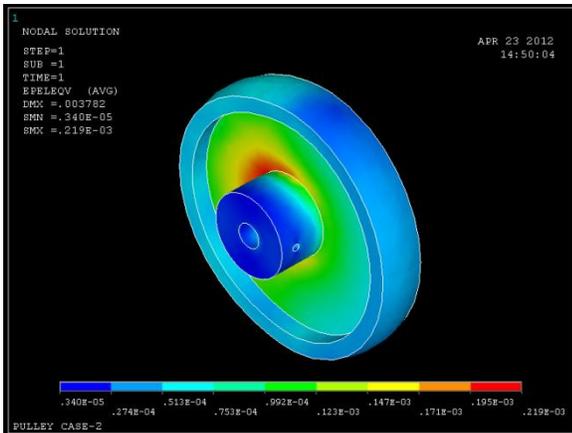


Fig. 8: Von Mises strain developed in Gray Cast Iron

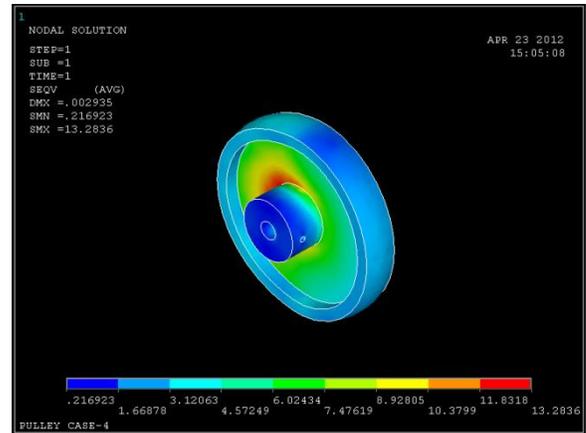


Fig. 11: Von Mises stress developed in Al alloy, 1460

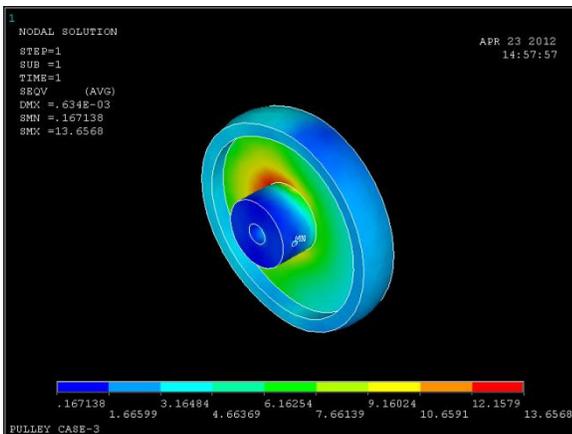


Fig. 9: Von Mises stress developed in Alumina

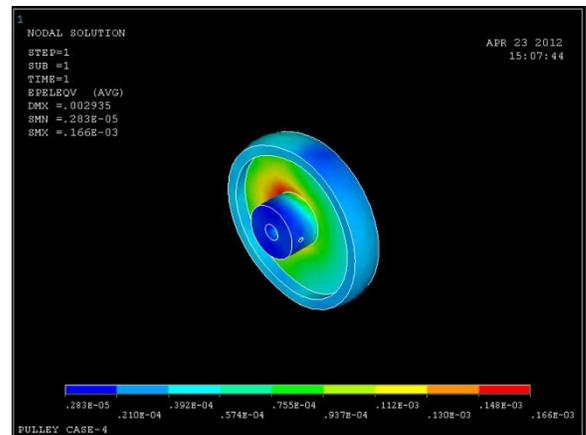


Fig. 12: Von Mises strain developed in Al alloy, 1460

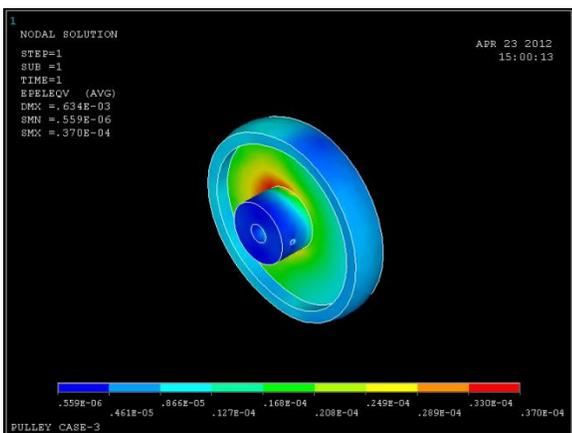


Fig. 10: Von Mises strain developed in Alumina

IV. RESULTS & DISCUSSIONS

Stress analysis of flat pulleys for various materials is depicted in Table 1. Stress and strain analysis is shown in Fig. 13 and Fig. 14 for various materials respectively.

Table 1 - Results of ANSYS for stress for various material of Pulley [N/mm²]

Sr. No.	Material	Stress	Strain
1	Polycarbonate Molded	12.971	0.00942
2	Gray Cast Iron	13.563	0.000219
3	Alumina	13.656	0.000037
4	Al alloy,1460 pressed profile	13.283	0.000166

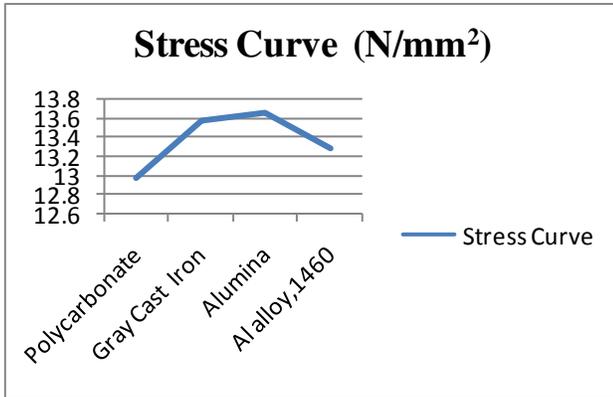
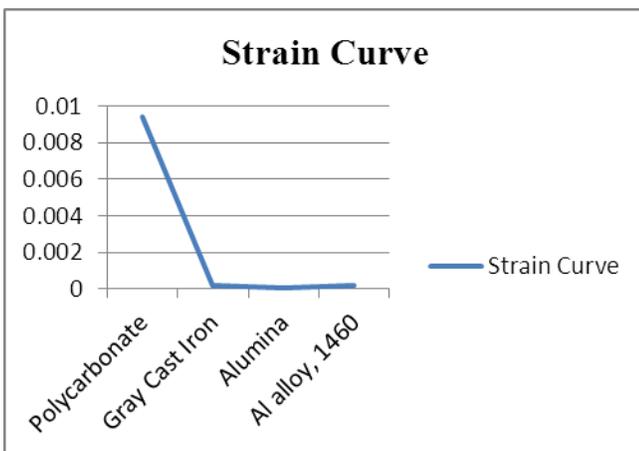


Table 3 - Comparison of FEA and Theoretical Stress Results [N/mm²]

Material	FEA Stress	Theoretical Stress	Variation %
Alumina	13.65	11.39	16.55



V. CONCLUSIONS

After the investigations, the real reason of the failure was determined to be associated with the material. It is also observed that, the maximum stress present in pulley varies for various material conditions but a little. It is seen that maximum Von Mises stress for four different pulleys are almost same and the effective factor of safety will be criterion for selecting material of pulleys. Strain doesn't have so much impact on the pulley, therefore, without considering strain we can concentrate on the factor of safety. It can be concluded that the ANSYS results with the assumption of load seems to be acceptable for all the 4 cases. Alumina can replace C.I. for pulley material which is light in weight, corrosion resistant and suitable for heavy duty applications. Variation in the results of FEA developed stress and Theoretical developed stress is 16.55% (approximately), which is in reasonable units.

REFERENCES

- [1] N.W.Pathan, Dr.D.V.Bhope & Prof.S.D Khamankar "Investigation of stress in Flat belt pulley by FEM and photo elasticity" *International Journal of Engineering Science and Technology (IJEST)*, 3(2011) 7444-7451.
- [2] Gregor Cepon,Lionel Manin & Miha Boltezar "Validation of a Flexible Multibody Belt-Drive Model" *Journal of Mechanical Engineering* , 57(2011) 539-546.
- [3] Simon Kulove, Leon Kos & Jozef Duhovnik "Mesh Smoothing with Global Optimization under Constraints" *Journal of Mechanical Engineering* 57(2011) 555-567.
- [4] Yilmaz D., Kursat H.& Ibrahim A."Finite element analysis of failure in rear mounted mower pulley" *Journal of food Agriculture and Environment* 7(2009) 856 – 868.
- [5] Ergin KILIC, Melik DOLEN & Ahmet KOKU, "Analysis and estimation of motion transmission errors of a timing belt drive" *Turk J.Elec. Eng. & Comp. Sci.* 18(2010) 906-944.
- [6] Serge Abrate "Vibrations of belts and belt drives." *Mechanism& Machine Theory*, 27 (1992) 645-659.

$$FOS = \frac{\text{Yield Stress of Pulley Material}}{\text{Developed Stress in Pulley Material}} \dots\dots(1)$$

Factor of safety was calculated for all the four materials and Alumina was found to be suitable for heavy load applications (Table 2). Al alloy, 1460 is having highest factor of safety but due to its higher cost it may not be substitute of Gray Cast Iron.

Table 2 – Calculated factor of safety for various materials

Sr. No.	Material	Factor of Safety
1.	Polycarbonate Molded	1.99
2.	Gray Cast Iron	3.38
3.	Alumina	7.69
4.	Al alloy, 1460	25.9

Theoretical developed stress was determined from the load and net area in contact with the pulley. Table 3 shows comparison of results.

COMPOSITIONAL ANALYSIS OF PbPc and CuPc THIN

S. Sivamalar*¹, J. Shanthi², Pon Kalugasalam¹

¹Department of Physics, Tamilnadu College of Engineering, India

²Department of Physics, Avinashilingam University, India

Abstract: The thin film of Lead Phthalocyanine (PbPc) and Copper Phthalocyanine (CuPc) on glass are prepared by Vacuum deposition method. Deposition of PbPc and CuPc on pre-cleaned glass substrates under the pressure of 10^{-6} Torr are achieved by slowly varying the current. The rate of evaporation is properly controlled and maintained constant during all the evaporations. The thicknesses of the films are 150 nm, 300 nm and 450 nm on glass substrate. The thickness of sample 450 nm annealed at 323 K and 373 K temperature. The sample has been analysed by Energy Dispersive Analysis using X-Rays (EDAX), Fourier Transform Infrared spectroscopy (FTIR) in order to get compositional analysis of the PbPc and CuPc thin film. The samples were evaluated by EDAX technique. The sample 450 nm is evaluated by Fourier Transform Infrared spectroscopy. The spectral pattern in $1000 - 1200 \text{ cm}^{-1}$ region depends strongly on the molecular structure of the complexes and its chemical structure for the central metal.

Keywords: Phthalocyanine, EDAX technique and Fourier Transform Infrared spectroscopy.

I. INTRODUCTION

The search for materials suitable for low cost, versatile electronic devices has stimulated interest in organic thin film transistors [OTFTs] and sensors. It has led in recent years to an extensive investigation of a range of metal substituted phthalocyanines [1,2]. Application of OTFTs as chemical sensors has shown promise in the development of electronic noses and in nerve agent detection [3-5]. A key issue regarding the widespread production of OTFTs is the long term stability and device integrity in ambient operating conditions [6,7]. Among the small molecule based OTFTs, pentacene OTFTs have received significant attention regarding instability to ambient components such as oxygen and humidity [8-12]. Several mechanisms have been proposed to explain this instability in pentacene OTFTs, including water adsorption in grain boundaries [10,11] and oxygen generated impurities[13].

Phthalocyanines have potential applications in optical logic display devices, electrophotography, security printing, gas detectors [14], solar cells [15, 16], sensitizers and colour filters [17]. These materials are generally p-type semiconductors and have the advantage of being sufficiently stable towards chemicals and heat. They can be easily sublimed, resulting in high purity thin films without decomposition. The physicochemical properties can be altered by changing the metal ion. Film properties of this prototype organic semiconductor are dependent on the evaporation rate, substrate temperature and post-evaporation annealing [17, 18]. Photovoltaic devices made from organic pigments have reached power conversion efficiency of a few percent [19, 20] that is much lower than those of their inorganic combinations.

This paper deals with the structural and compositional studies of lead phthalocyanine and Copper phthalocyanine thin films prepared by vacuum evaporation technique. These materials are an organic semiconducting material which has various functional groups. So this paper deals with the Fourier transform infrared spectroscopy of functional group analysis.

II. EXPERIMENTS

The powder of PbPc (80% dye, Sigma Aldrich company, Bangalore, India) is kept in a molybdenum boat (100 A current rating) heated with high current controlled by a transformer. The transformer is capable of supplying 150 amps at 20 volts which is used to provide the accessory current for heating the molybdenum source. It is used for the evaporation process. Prior to evaporation, the evaporant material is carefully degassed at lower temperature for about 45 minutes with the closed shutter. Thin films of PbPc are deposited at room temperature on pre-cleaned glass substrates under the pressure of 10^{-6} Torr using a (12 A 4D Hind Hivac, India) coating unit. The rate of evaporation is properly controlled and maintained constant during all the evaporations. Rotary drive is employed to maintain uniformity in film thickness. The thicknesses of the films are 150 nm, 300 nm and 450 nm. The thickness of the films is measured by Quartz crystal monitor. This procedure is used for preparing PbPc and CuPc thin film on pre-cleaned glass substrate. The adhesion of the films to the substrate seems to be extremely good. The samples prepared in a similar environment are used for studying their various properties. The structures and composition of the films are examined by EDAX and FTIR.

III. RESULTS AND DISCUSSION

III.1. EDX Analysis

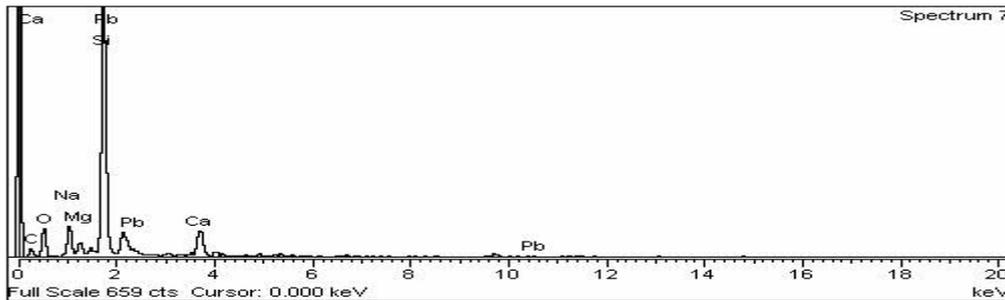
Energy dispersive analysis using X- ray (EDAX) helps us to determine the elemental contents in the specimen very accurately. A core level electron of a surface atom of PbPc is removed by an impinging electron or an X-ray photon. The excited atom decays to a lower energy state through an electronic rearrangement in which an additional electron from a higher level is knocked out leaving the atom in a doubly ionized state. The energy difference between the two states is given to the ejected electron, which will have kinetic energy characteristic of the atom.

III.1.1. Effect of thickness

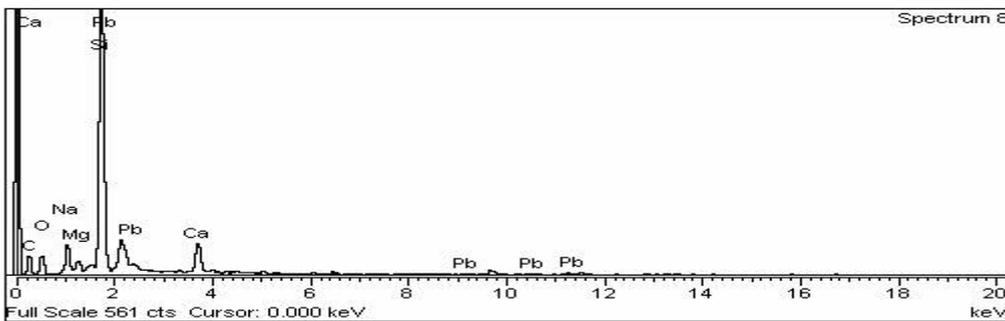
The elemental composition of the PbPc and CuPc films deposited at different thickness and annealed temperature were evaluated by EDAX technique. The Figure 1 & 2 (a, b and c) shows the EDAX spectrum of the PbPc and CuPc films respectively. The percentage of lead and copper increases with thickness in both PbPc and CuPc films respectively (0.77, 1.30 and 6.18 for 150, 300 and 450 nm of PbPc & 0.89, 1.93 and 8.78 for 150, 300 and 450 nm of CuPc respectively). The percentage of carbon increases with thickness (30.93, 50.47 and 62.00 for 150, 300 and 450 nm of PbPc & 38.34, 49.25 and 60.82 for 150, 300 and 450 nm of CuPc respectively). The nitrogen amount is very poor in both films. The film deposited at 303 K with different thickness has a uniform composition profile. However, the films deposited at 303 K with different thickness are required for the homogenization.

The peak corresponding to carbon, silicon, sodium and oxygen shows a transition from L to the K-shell which can be termed as a K-Alpha peak ($K\alpha$). The peaks corresponding to lead and Copper shows three types of transitions M-Alpha ($M\alpha$), M-beta ($M\beta$), M-Gamma ($M\gamma$). These all transitions are shown in the EDAX patterns of PbPc and CuPc thin films of different thickness.

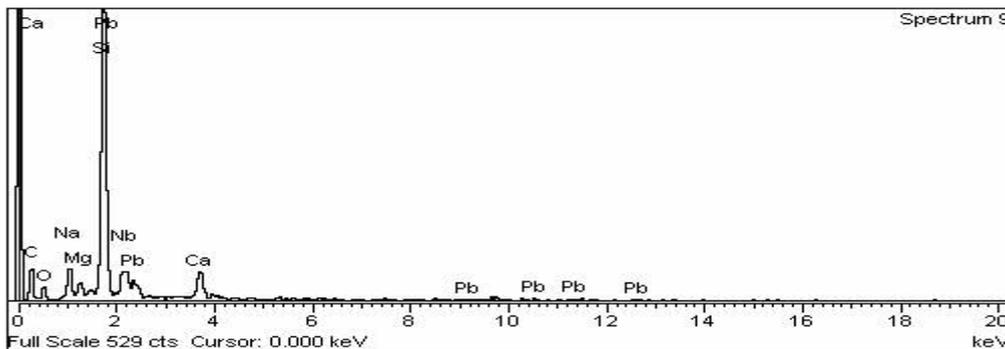
Figure 1. EDAX Patterns of PbPc Films for thicknesses



(a) 150 nm

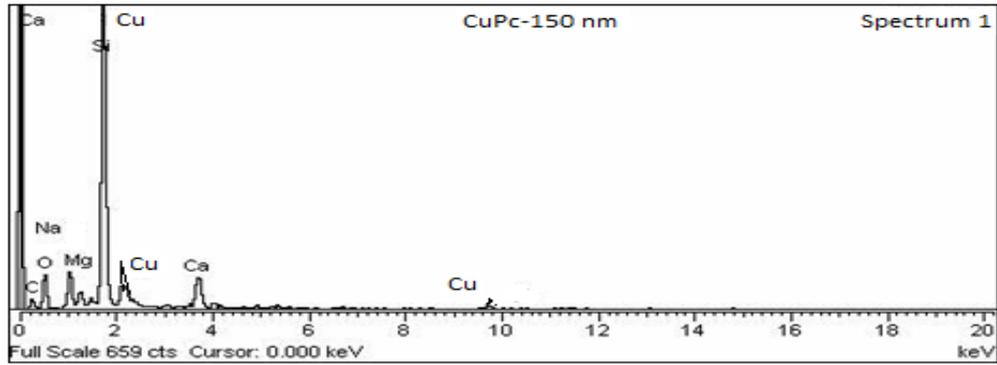


(b) 300 nm

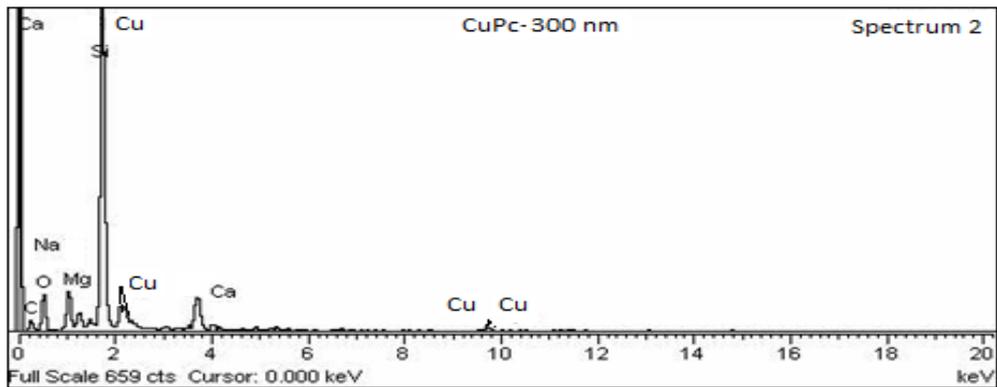


(c) 450 nm

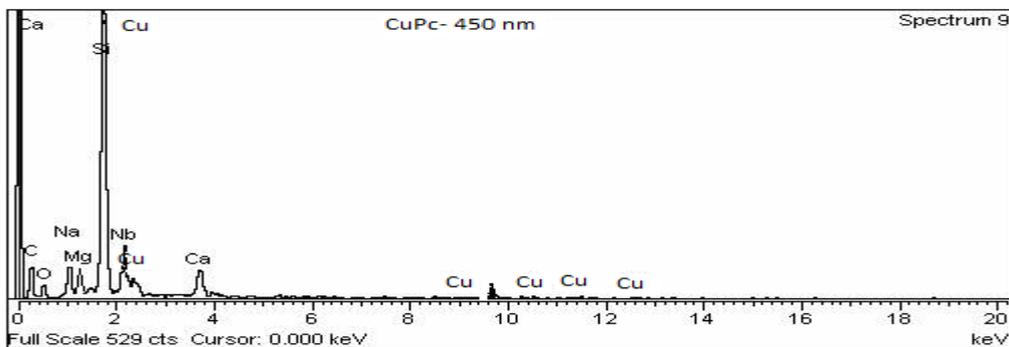
Figure 2. EDAX Patterns of CuPc Films for thicknesses



(a) 150 nm



(b) 300 nm



(c) 450 nm

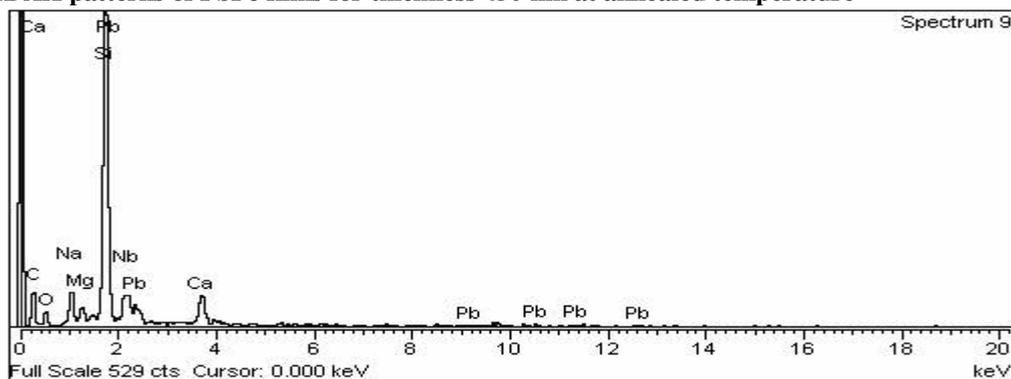
III.1.2. Effect of temperature

Figure 3 & 4 (a, b and c) shows the EDAX patterns of PbPc and CuPc thin films of thickness 450 nm at room temperature (303 K) and annealed at 323 K and 373 K temperature. The percentage of lead and Copper are same in both cases and its amount is low. The percentage of carbon is increased from 62 to 84.81 for Lead and from 60 to 80 for Copper when annealed at high temperature (373 K). The nitrogen amount is very poor. Generally elements of low atomic number are difficult to detect by EDAX. So, in the cases of PbPc and CuPc films hydrogen is difficult to detect.

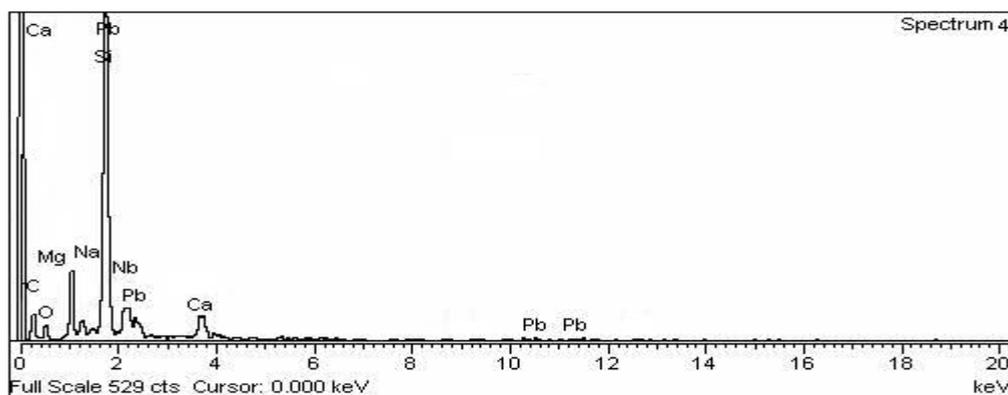
The presence of Silicon, Sodium and Oxygen peaks are due to the glass substrates [21]. The composition of the films deposited at higher thickness and annealed temperatures are found to be more in percentage. It is observed that the deposited annealed films of 373 K slightly deviation from stoichiometry and the film deposited at 303 K has a uniform composition profile. However, the film deposited at 303 K is required for the homogenization. The films deposited and annealed at 373 K only, the thermodynamically stable phase are present and it exhibited a microcrystalline structure with a uniform composition at the micrometer. From this analysis, it is found that the film possess PbPc and CuPc stoichiometry. The small deviation for the stoichiometry may be attributing to surface oxidation.

The peak corresponding to carbon, nitrogen, silicon, sodium and oxygen shows a transition from L to the K-shell which can be termed as a K-Alpha peak ($K\alpha$). The peaks corresponding to lead and Copper shows three types of transitions – M-Alpha ($M\alpha$), M-beta ($M\beta$), M-Gamma ($M\gamma$). All these transitions are shown in the EDAX patterns of PbPc and CuPc thin films.

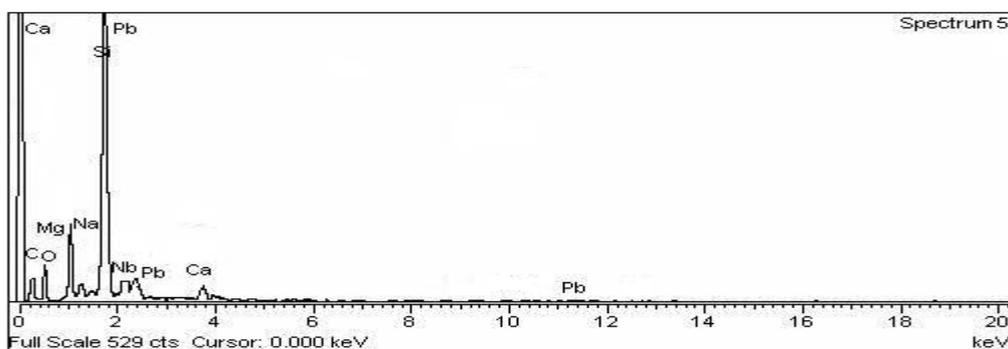
Figure 3 EDAX patterns of PbPc films for thickness 450 nm at annealed temperature



(a) 303 K

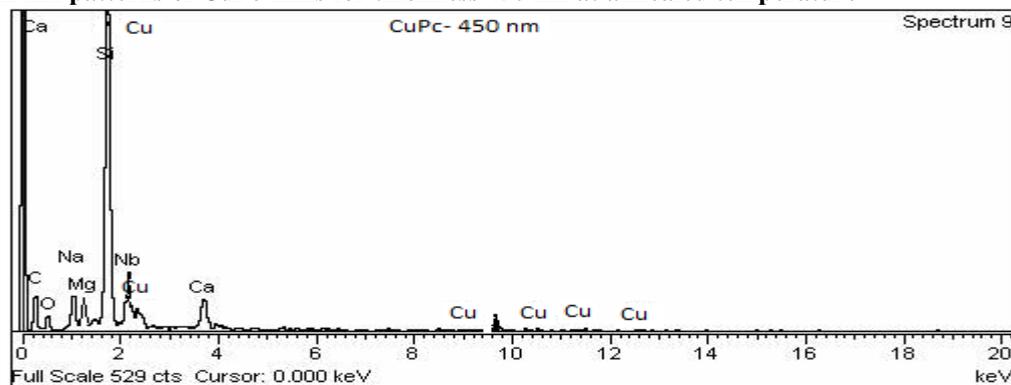


(b) 323 K

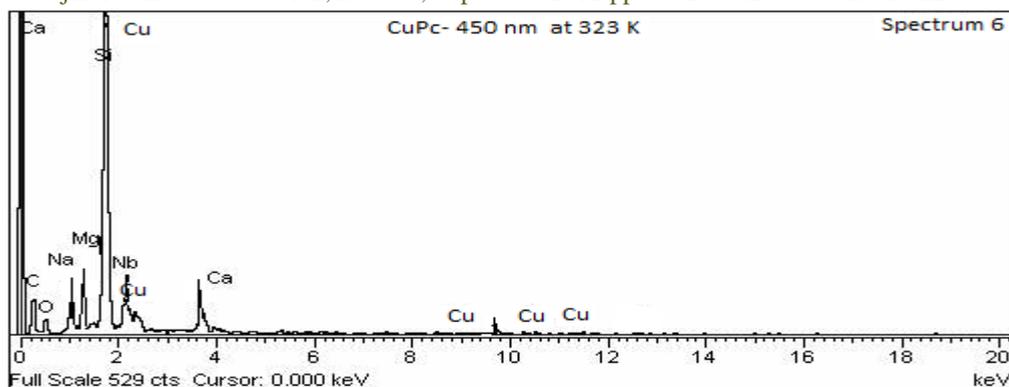


(c) 373 K

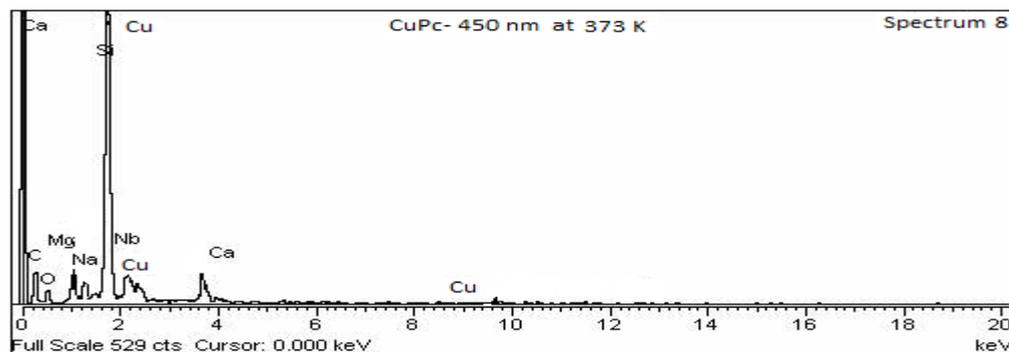
Figure 4 EDAX patterns of CuPc films for thickness 450 nm at annealed temperature



(a) 303 K



(b) 323 K



(c) 373 K

III.2. FT-IR Analysis

Figures 5 to 14 shows the infrared spectrum of PbPc and CuPc thin film of 150 nm, 300 nm and 450 nm thickness prepared at room temperature and 450 nm annealed at 323 K and 373 K temperature. The main peak at $725.23 \pm 4 \text{ cm}^{-1}$ is attributed to non-planar deformation vibrations of the C–H bonds of benzene rings [22, 23] and the C–H in plane bonding at $1112.93 \pm 4 \text{ cm}^{-1}$ [22, 24]. A medium band at $771.53 \pm 4 \text{ cm}^{-1}$ also corresponds to non-planar vibrations (Out-of-plane-bonding) of the C–H bonds [22, 25]. The following assignments can be made for the other ligand bands, the $3049.46 \pm 4 \text{ cm}^{-1}$ band to aromatic C–H symmetric stretching vibrations and the 1606.70 ± 4 and $1479.40 \pm 4 \text{ cm}^{-1}$ band to C–C benzene ring skeletal stretching vibrations [22, 26, 27]. So it is expected to be intense compared to the C–N stretching at $1078.21 \pm 4 \text{ cm}^{-1}$.

One strong intensity band at $881.47 \pm 4 \text{ cm}^{-1}$ indicates the extraordinary stability of the metal phthalocyanines due to the strong bonding between the metal ion and the four surrounding nitrogen atoms in the pyrrole rings [22, 28]. The peaks in the $700\text{--}400 \text{ cm}^{-1}$ interval originate most probably from vibrations in the benzene ring in interaction with the pyrrole ring [22].

The band at $1332 \pm 6 \text{ cm}^{-1}$ is assigned to the C–C stretching in isoindole. The bands appearing at 882 ± 10 and $728 \pm 5 \text{ cm}^{-1}$ are assigned to the C–H bonding out of plane deformations. The bands appearing at 1284 ± 6 and $1162 \pm 4 \text{ cm}^{-1}$ is assigned to the C–N in isoindole and in plane band stretching vibration [29].

3.2.1 Effect of thickness

Figures 5 to 7 shows the referred spectrum of PbPc thin film of 150 nm, 300 nm and 450 nm respectively. The peaks are formed at 439.77 ± 4 , 495.71 ± 4 and $605.65 \pm 4 \text{ cm}^{-1}$ due to vibrates in the benzene ring in interaction with the byrrole ring. The intensity of the peak increases with increase in the thickness of the films. The strong intensity band at $881.47 \pm 4 \text{ cm}^{-1}$ decreases with increase in the thickness indicates the stability of the metal Phthalocyanines due to the strong bonding. The intensity of the peak at $3049.46 \pm 4 \text{ cm}^{-1}$ for aromatic C-H symmetric stretching vibration and peak at $2536.39 \pm 4 \text{ cm}^{-1}$ decreases with increase in the thickness. The peaks at $1606.70 \pm 4 \text{ cm}^{-1}$, $1479.40 \pm 4 \text{ cm}^{-1}$ and $1330.88 \pm 4 \text{ cm}^{-1}$ bands to C-C benzene ring skeletal stretching vibrations. The C-C and C-H benzene ring peak intensities increase with thickness. The intensity of C-N stretching at $1078.21 \pm 4 \text{ cm}^{-1}$ decreases with increase the thickness of the film. The absence of the band at $1006\text{--}1008 \text{ cm}^{-1}$ and $1539 \pm 4 \text{ cm}^{-1}$ suggest that the sample does not contain any metal free phthalocyanine. The peaks are at $1078.21 \pm 4 \text{ cm}^{-1}$, $1112.93 \pm 4 \text{ cm}^{-1}$ and $1159.22 \pm 4 \text{ cm}^{-1}$ depends on the molecular structure of the complexes and its chemical structure for the central metal.

Figures 8 to 10 shows the spectrum of CuPc thin film of 150 nm, 300 nm and 450 nm respectively. In a metal phthalocyanine, spectrum have a band at 1003 cm^{-1} originates from N–H bending vibration. This band is absent in metal phthalocyanine. This is because in copper phthalocyanine, metal is replaced by metal cation. Metal ligand vibration band is observed in both copper and lead phthalocyanine at 940 and 960 cm^{-1} . The peaks are formed at 495.85 ± 4 and $581.4571 \pm 4 \text{ cm}^{-1}$ due to vibrates in the benzene ring in interaction with the byrrole ring. The intensity of the peak increases with increase in the thickness of the films. The strong intensity band at $771.50 \pm 4 \text{ cm}^{-1}$ decreases with increase in the thickness indicates the stability of the metal Phthalocyanines due to the strong bonding. The intensity of the peak at $3038.90 \pm 4 \text{ cm}^{-1}$ for aromatic C-

H symmetric stretching vibration and peak at $2678.89 \pm 4 \text{ cm}^{-1}$ was decreases with increase the thickness of the film. The peaks at $1391.09 \pm 4 \text{ cm}^{-1}$, $1490.00 \pm 4 \text{ cm}^{-1}$ and $1900.25 \pm 4 \text{ cm}^{-1}$ bands to C-C benzene ring skeletal stretching vibrations. The C-C and C-H benzene ring peaks intensities increases with thickness. The intensity of C-N stretching at $1091.34 \pm 4 \text{ cm}^{-1}$ was decreases with increase the thickness of the film. The absence of the band at $1006-1008 \text{ cm}^{-1}$ and $1539 \pm 4 \text{ cm}^{-1}$ suggests that the sample does not contain any metal free phthalocyanine. The peaks are at $1091.34 \pm 4 \text{ cm}^{-1}$, $1391.09 \pm 4 \text{ cm}^{-1}$ and $1491.22 \pm 4 \text{ cm}^{-1}$ depends on the molecular structure of the complexes and its chemical structure for the central metal.

The presence of bands at 712 and 739 cm^{-1} suggests the presence of α -phase and the presence of bands at 730 and 752 cm^{-1} suggests that β -phase is also present in CuPc. The FTIR spectroscopic analysis of PbPc and CuPc suggest that these metal phthalocyanines contain both α - and β -phases. The presence of band at 771 cm^{-1} in CuPc spectrum suggests presence of β phase.

Figure 5 FTIR Spectrum of PbPc film of thickness 150 nm prepared at room temperature

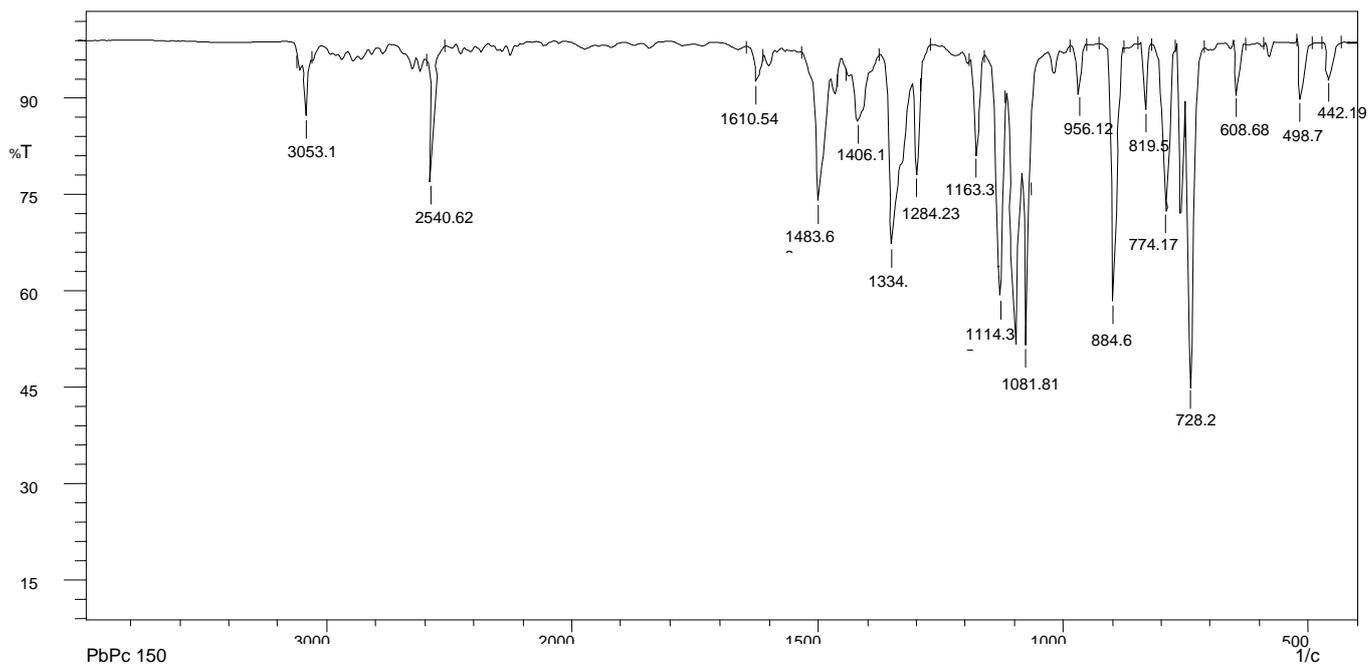
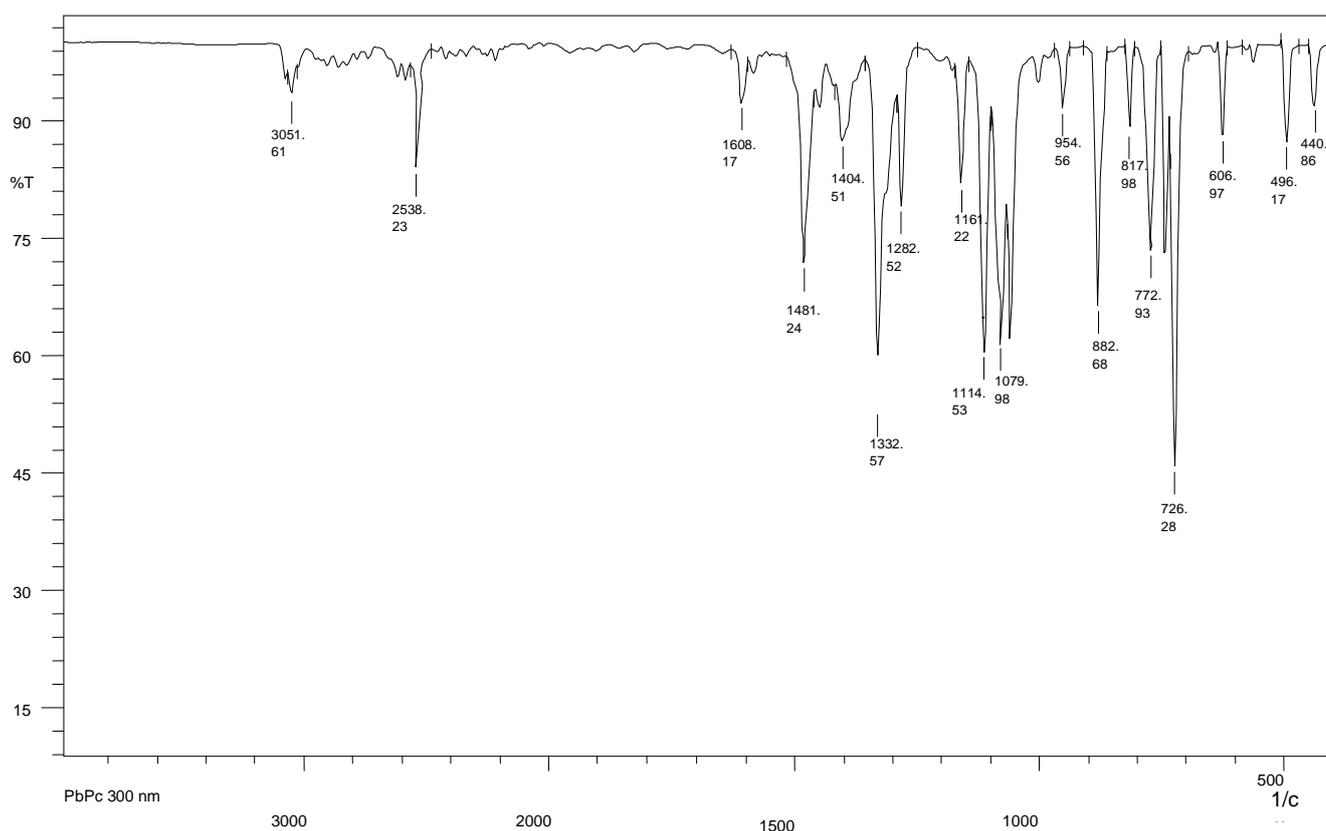


Figure 6. FTIR Spectrum of PbPc film of thickness 300 nm prepared at room temperature



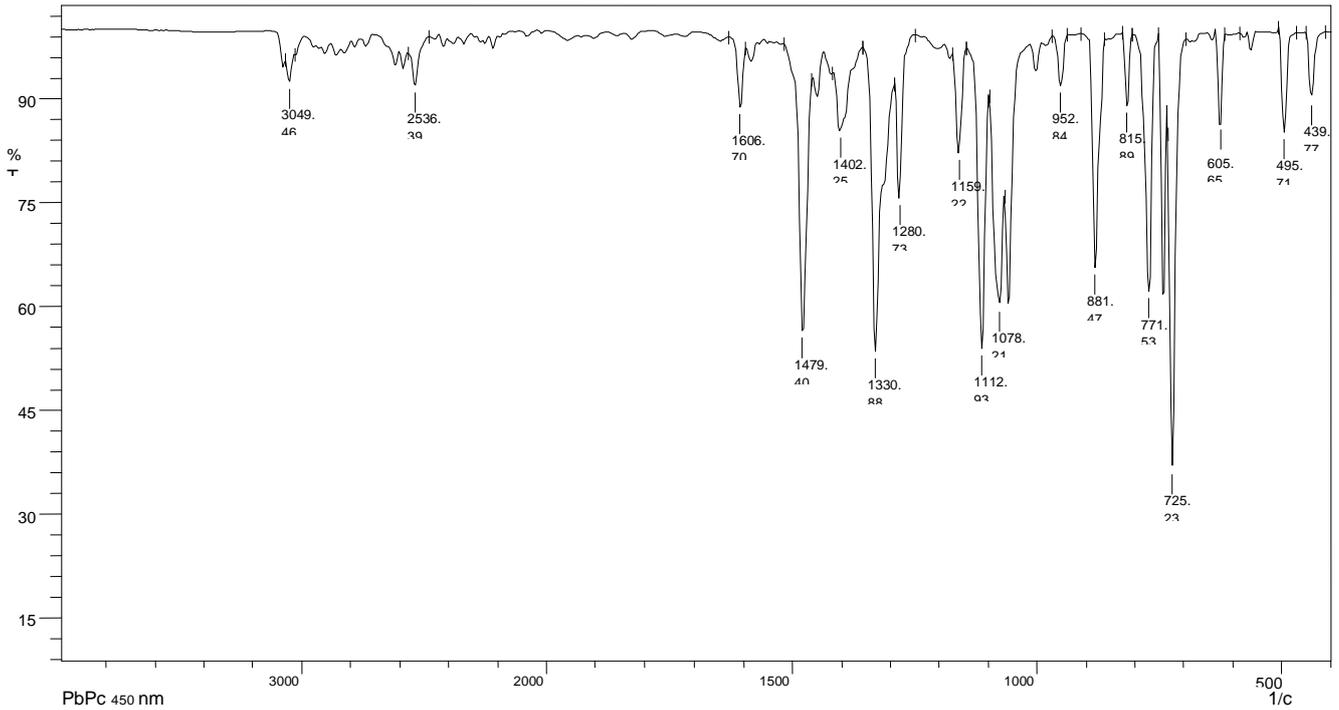


Figure 8 FTIR Spectrum of CuPc film of thickness 150 nm prepared at room temperature

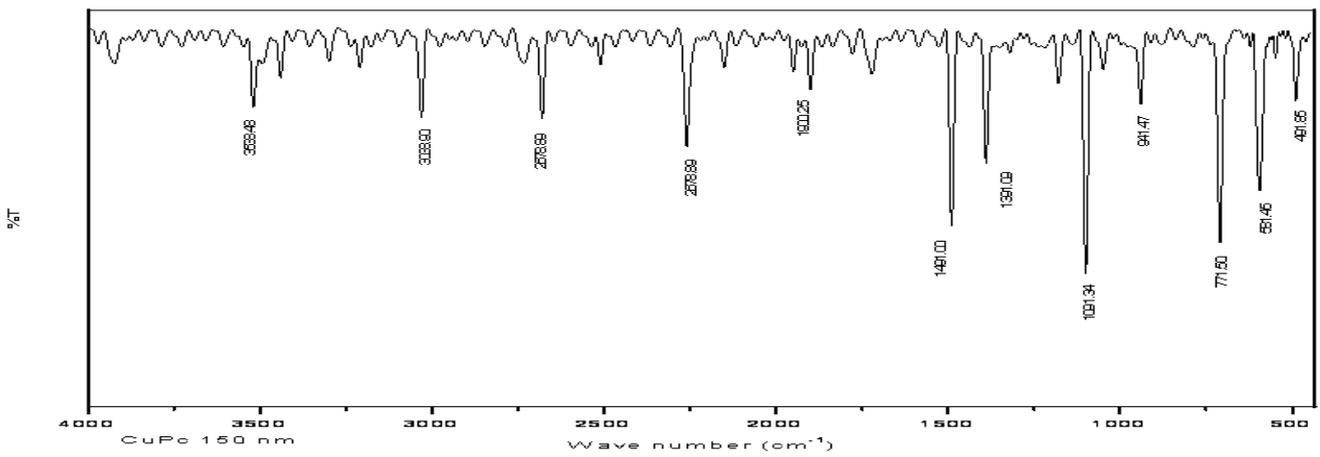


Figure 9 FTIR Spectrum of CuPc film of thickness 300 nm prepared at room temperature

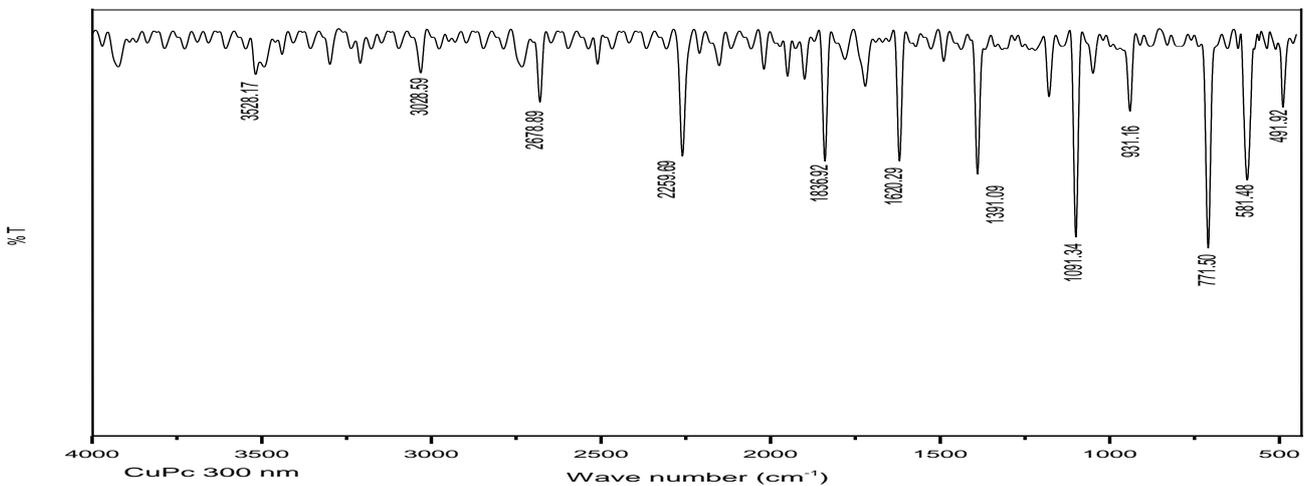
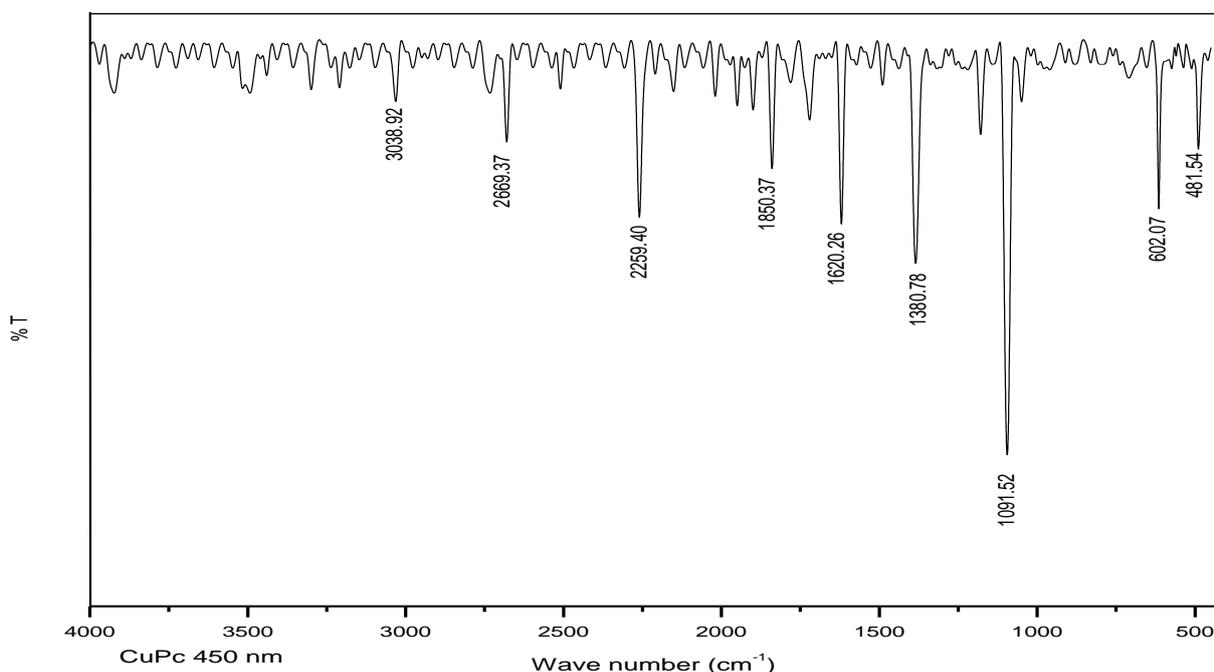


Figure 10: FTIR Spectrum of CuPc film of thickness 450 nm prepared at room temperature.



The Reflectance (R) / Transmission (T) ratios for drop casting films over glass substrate are shown in Table 1. The R/T is found to be around 1.0 in case of PbPc films, which suggests that in drop casting PbPc films have their molecular plane oriented closer to parallel to plane of the glass substrate. In case of CuPc drop casting films, R/T value is around 1.8, which suggests that CuPc molecules have similar preferential orientation as PbPc drop casting films.

Table 2 Comparison of reflectance spectra of films with transmission

MPC	Reflectance (R)	Transmission (T)	R/T
PbPc	0.78	0.79	1.0
CuPc	2.42	1.34	1.8

III.2.2 Effect of temperature

The FT-IR patterns of PbPc and CuPc thin films for thickness 450 nm at different annealed temperatures are shown in the Figures 11 to 14. The benzene ring peaks in the 700 - 400 cm^{-1} interval have increased the intensity with increase in the temperature. The main peaks at $725.23 \pm 4 \text{ cm}^{-1}$ and $771.53 \pm 4 \text{ cm}^{-1}$ attributed to non-planar deformation vibrations of the C-H bonds. There is no change in this peak due to annealing. The strong bonding between the metal ion and the four surrounding nitrogen atom in the pyrrole ring peak at $881.47 \pm 4 \text{ cm}^{-1}$ indicates increasing intensity with annealing temperature. The C-H in-plane bonding at $1112.93 \pm 4 \text{ cm}^{-1}$ peaks intensity increase with temperature. The spectral pattern in this region 1200-1000 cm^{-1} depends strongly on the molecular structure of the complexes and its chemical structure for the central metal. The middle peak originates from the vibration mode of a pyrrole ring and other two peaks are assigned to the in-plane deformation vibration of C-H bonding in the ring. The middle band has the highest intensity among the three peaks in the usual MPC's with D_4 molecular symmetry; however, the band shows lowest intensity in to the case of PbPC and CuPc. This should be attributed to the molecular structure of PbPc and CuPc [29, 30, 31]. Due increase the temperature, the absence of the band at 3049.46 ± 4 and $2536.39 \pm 4 \text{ cm}^{-1}$ in the phthalocyanine complexes results in the replacement of hydrogen by metal cation [29].

Figure 11 FTIR Spectrum of PbPc film of thickness 450 nm annealed at 323 K

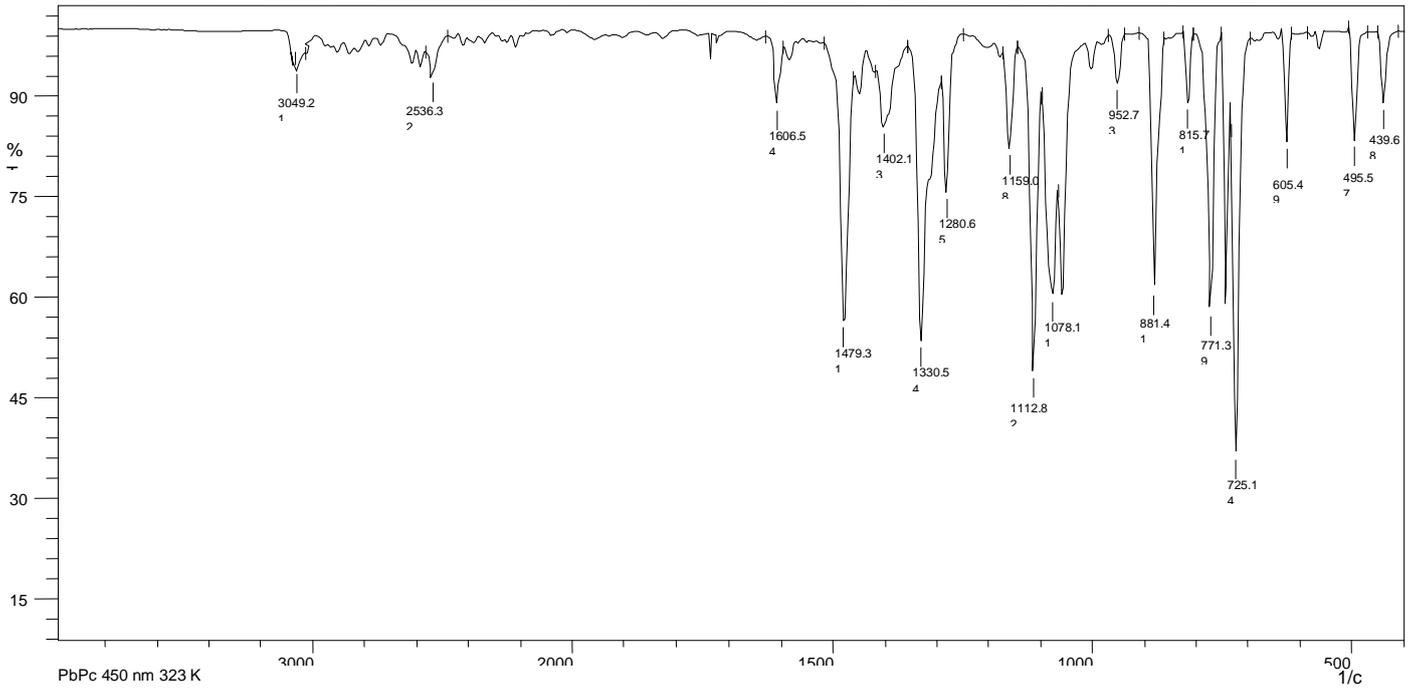


Figure 12 FTIR Spectrum of PbPc film of thickness 450 nm annealed at 373 K

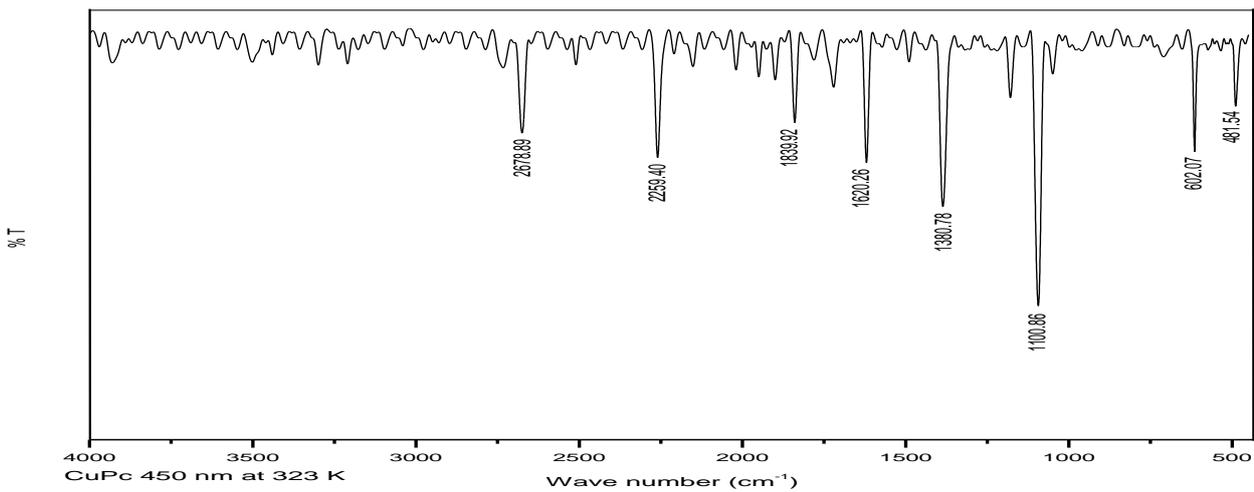


Figure 13 FTIR Spectrum of CuPc film of thickness 450 nm annealed at 323 K

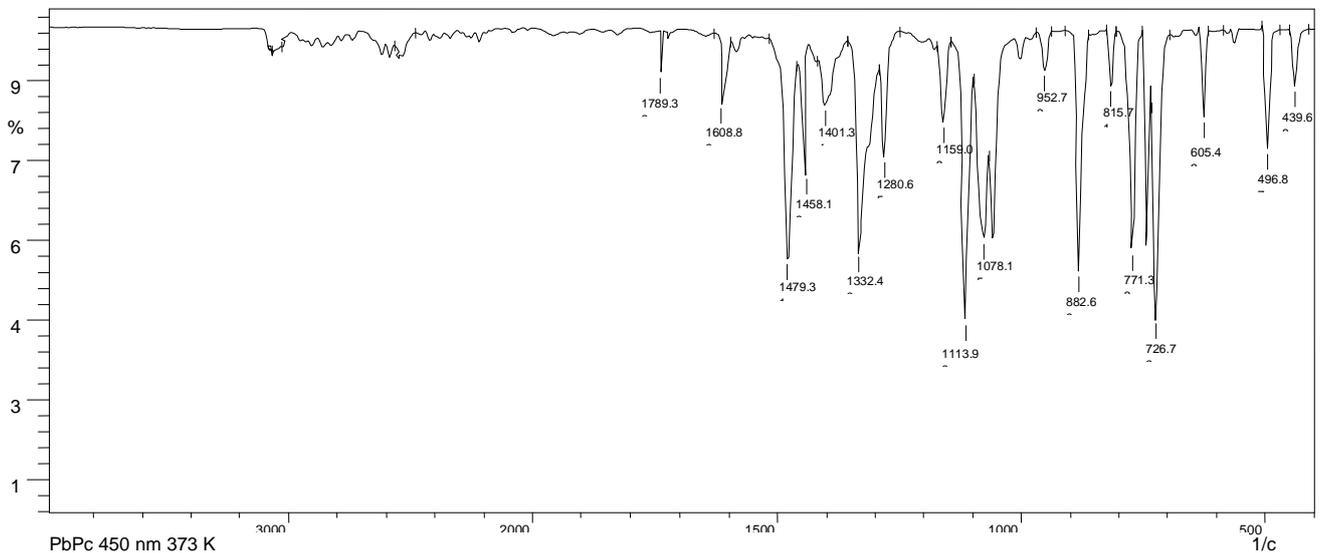
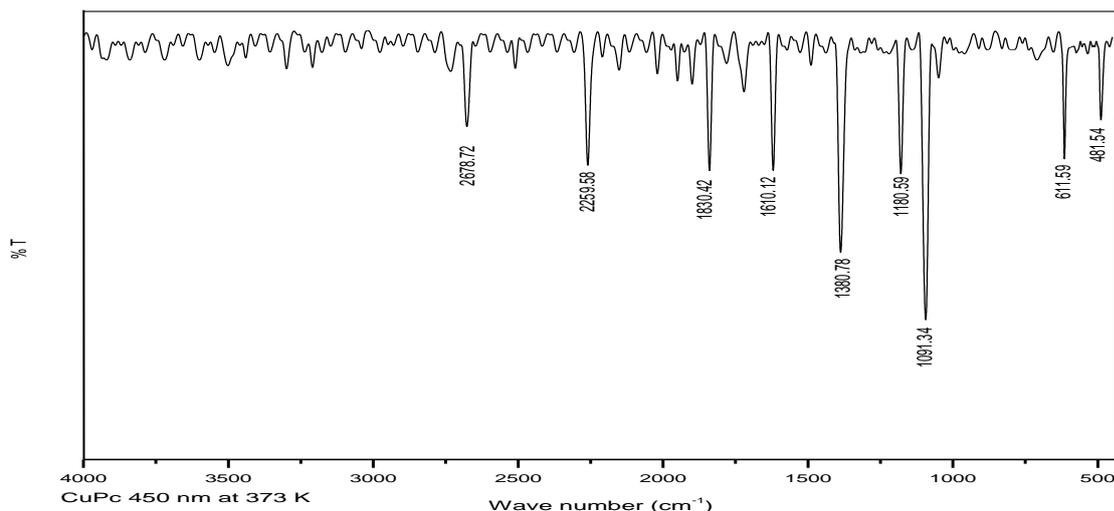


Figure 14 FTIR Spectrum of CuPc film of thickness 450 nm annealed at 373

The material characterization of these films by FT-IR spectroscopy clearly provides a convenient diagnostic technique in the development of PbPc and CuPc based thin film sensors [29].

IV. CONCLUSION

EDAX analysis is used to find the elemental composition. The percentage of lead, copper and carbon were increases with thickness. The presence of Silicon, Sodium and Oxygen peaks are due to the glass substrates. The percentage of carbon is increased when annealed at high temperature (373 K). The composition of the films deposited at higher thickness and annealed temperatures are found to be in more percentage.

FT-IR analysis identifies the presence of functional groups which absorb at definite frequencies. The spectral pattern in this region depends strongly on the molecular structure of the complexes and its chemical structure for the central metal. The FTIR spectroscopy analysis indicates presence of both α and β -phases of phthalocyanines in both the samples. We have utilized grazing angle FTIR spectroscopy to analyze molecular orientations in phthalocyanine films deposited over glass substrate. The CuPc samples showed preferential orientation along parallel direction of the substrate surface in vacuum coated thin films. Our results suggest that metal ions play an important role in defining molecular orientation in thin films.

REFERENCES

1. C. D. Dimitrakopoulos and P. R. L. Malenfant, *Adv. Mater.* (Weinheim, Ger.) 14, 99(2002).
2. G. Horowitz, *Adv. Mater.* (Weinheim, Ger.) 10, 365 (1998).
3. R. D. Yang, J. Park, C. N. Colesniuc, I. K. Schuller, W. C. Trogler, and A.C. Kummel, "Ultralow drift in organic thin-film transistor chemical sensors by pulsed gating" *J. Appl. Phys.* 102, 034515, (2007).
4. F. Liao, C. Chen, and V. Subramanian, "Organic TFTs as gas sensors for electronic nose applications", *Sensors and Actuators B-Chemical*. Vol. 107, pp. 849-855, (2005).
5. J. B. Chang, V. Liu, V. Subramanian, K. Sivula, C. Luscombe, A. Murphy, J. Liu, and J. M. J. Frechet. "Printable polythiophene gas sensor array for low-cost electronic noses" *Journal of Applied Physics*, 100(1):014506, (2006).
6. R. Ben Chaabane, A. Ltaief, L. Kaabi, H. Ben Ouada, N. Jaffrezic-Renault and J. Davenas, "Influence of ambient atmosphere on the electrical properties of organic thin film transistors" *Mater. Sci. Eng., C* 26, 514, (2006).
7. G. Guillaud, J. Simon, and J. P. Germain, *Metallophthalocyanines: "Gas sensors, resistors and field effect transistors"* *Coord. Chem. Rev.* 178-180, 1433 -1484, (1998).
8. S. Cipolloni, L. Mariucci, A. Valletta, D. Simeone, F. De Angelis and G. Fortunato "Aging effects in high performance Pentacene Thin Film Transistors." *Thin Solid Films*, vol. 515 , pp. 7546-7550 (2007).
9. D. Knipp, T. Muck, A. Benor, and V. Wagner, "Environmental stability and electronic transport of pentacene thin film transistors" *J. Non-Cryst. Solids* 352, 1774 (2006).
10. Y. Qiu, Y. Hu, G. Dong, L. Wang, J. Xie, and Y. Ma, "H₂O effect on the stability of organic thin-film field-effect transistors" *Appl. Phys. Lett.* 83,1644 (2003).
11. Z.-T. Zhu, J. T. Mason, R. Dieckmann, and G. G. Malliaras, "Humidity sensors based on pentacene thin-film transistors" *Appl. Phys. Lett.* 81, 4643 (2002).
12. C. R. Kagan, A. Afzali, and T. O. Graham, "Operational and environmental stability of pentacene thin-film transistors" *Appl. Phys. Lett.* 86, 193505 (2005).
13. Y. Natsume, "Characterization of solution-processed pentacene thin film transistors" *Phys. Status Solidi A* 205, 2958-2965 (2008).
14. Mrwa A., Friedrich M., Hofmann A. and Zahn D.R.T. (1995) 'Response of PbPc to high NO₂ concentration', *Sensors and Actuators B*, Vol. 24-25, pp. 596-599, (1995).
15. M. Pope, C.E. Swenberg, *Electronic Process in Organic Crystals*, Clarendon Press, Oxford, (1992).

16. Tang C.W. and Van Slyke S.A. 'Organic electroluminescent diodes', Appl. Phys. Lett., Vol. 51, p. 913 (1987).
17. Collin R.A. and Belgachi A. 'Electrical and optical studies on thin films of indium phthalocyanine chloride', Mater. Lett., Vol. 9, pp. 340-349, (1989).
18. Machida Y., Saito Y., Taomoto A., Nichogi K., Waragi K. and Asakawa S. 'Electrical Switching in Evaporated PbPc Films', Jpn. J. Appl. Phys. Vol. 28, pp 297-298, (1989).
19. B.A. Gregg, 'Excitonic Solar Cells' J. Phys. Chem. B 107, 4688 (2003).
20. Peumans P. and Forrest S.R. 'Very-high-efficiency double-heterostructure copper phthalocyanine/C₆₀ photovoltaic cells', Appl. Phys. Lett., Vol. 79, p. 126(2001).
21. Prabahar S. and Dhanam M. 'CdS thin films from two different chemical baths-structural and optical analysis', J. Crys. Growth, Vol. 285, pp. 41- 48. (2005).
22. A. Ahmad and R. A. Collins, "FTIR characterization of triclinic lead phthalocyanine" J. Phys. D: Appl. Phys. 24 1894-1897 (1991).
23. A.N. Sidorov, "Fourier-transform infrared and UV-vis spectroscopies of nickel phthalocyanine thin films", Opt. Spectrosc. 40,280,(1976).
24. A.J. Dann, H. Hoshi and Y. Maruyama, "Spectral dependence of the anisotropy of $\chi(3)$ of epitaxially grown vanadyl phthalocyanine film", J. Appl. Phys. 67, 1371,(1990).
25. G.N. Meshkova, A.T. Vartanyan and A.N. Sidory, "Fourier-transform infrared and UV-vis spectroscopies of nickel phthalocyanine thin films", Opt. Spectrosc. 43, 151(1977).
26. D. Campbell and R.A. Collins, "A study of the interaction between nitrogen dioxide and lead phthalocyanine using electrical conduction and optical absorption" Thin Solid Films 295, 277-282(1997).
27. W.J. Kroenke and M.E. Kenny, "Low Hue Photobleaches", Inorg. Chem. 3277-282,(1964).
28. T.Kobayashi, F. Kurokawa, N. Uyeda and E. Suito, "Structural Properties and UV to NIR Absorption Spectra of Metal-free Phthalocyanine (H2PC) Thin Films", Spectrochim. Acta. A 26, 1305,(1970).
29. R. Seoudi, G.S. El-Bahy and Z.A. El-Sayed, "FTIR, TGA and DC electrical conductivity studies of phthalocyanine and its complexes" J. Mole. Stru. 753, 119-126,(2005).
30. F.H. Moser and A.L. Thomas, "Phthalocyanine compounds", Reinhold publishing New York (1963).
31. Janczak and R. Kubiak, "Catalytic trimerization of 2- and 4-cyanopyridine isomers to the triazine derivatives in presence of magnesium phthalocyanine", Polyhedron 21, 265,(2002)

Channel Assignment Mechanism for Wireless Local Area Networks (WLANs) - A Survey

S. Gopalakrishnan, Dr. P. GaneshKumar,

(Assistant Professor, Department of ECE, PSNA college of Engineering and Technology, Dindigul, Tamil Nadu, India)
(Professor, Department of Information Technology, PSNA College of Engineering and Technology, Dinigul, Tamil Nadu, India)

Abstract: Wireless Local Area Networks (WLANs) are subjected to interference because of their working in Unlicensed Spectrum. Coverage ranges of WLANs are 30 to 300m. Channel Assignment is one of the most important problems in the WLANs. In this paper we present a Survey on the Channel Assignment techniques in IEEE 802.11. After analyzing different methods a qualitative comparison has been carried out for various methods by considering execution response, complexity and scalability. We then accomplish the Survey with many research descendants open for further scrutiny.

Keywords: Channel allocation, IEEE 802.11, wireless Networks, Channel Assignment.

I. Introduction

FOR an IEEE 802.11 WLAN with a only AP, there are many research efforts in the modification of MAC protocols to find an ideal transmission probability for an effective station to reach the maximum protocol capacity, granted that the number of effective stations associated with the AP is given [1]. The sensational evolution of the wireless/mobile user populace, coupled with the bandwidth requirements of Audio/video applications, requires efficient reuse of the inadequate radio spectrum allocated to wireless/mobile technology. A

total radio spectrum is to be divided into a set of disjointed channels that can be used simultaneously while minimizing interference in adjacent channel by allocating channels appropriately.

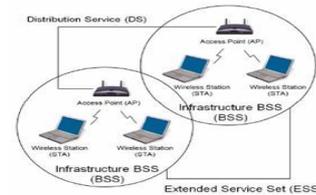


Fig. 1: Infrastructure based IEEE 802.11 WLANs

Channel allocation schemes can be divided in general into Fixed Channel Allocation schemes, Dynamic Channel Allocation schemes, and Hybrid Channel Allocation schemes, combining both FCA and DCA techniques. DCA schemes can be centralized or distributed. The centralized DCA scheme involves a single controller selecting a channel for each cell. The distributed DCA scheme involves a number of controllers scattered across the network (MSCs).

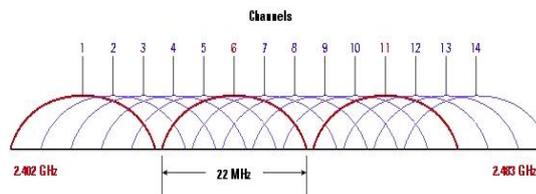


Fig 2: IEEE 802.11 Channel Allocations

II. System Under Consideration

2.1, Network topology

This survey focuses only on an IEEE 802.11 WLANs with an infrastructure network topology as shown in Fig 1, where APs and client resort to existing communication infrastructures such as legacy LAN to facilitate their communication. All communication accomplishment must be facilitated via this AP. This survey focuses only on an IEEE 802.11 WLAN with an infrastructure network topology as shown in Fig. 1, where APs and clients.

2.2, IEEE 802.11 Channels

The 802.11a specification today specifies 4 channels for the UNII1 band, 4 channels for the UNII@ band, and 4 channels for the UNII3 band. These channels are spaced at 20MHz apart and are considered non-interfering, however they do have a slight overlap in frequency spectrum. It is possible to use adjacent channels in adjacent cell coverage, but it is recommended when possible to separate adjacent cell channels by at least 1 channel. Figure 2 shows the channel scheme for the 802.11 bands.

III. Channel Assignment Schemes

3.1, GDCA Algorithm

GDCA algorithm focuses on Channel Assignment (CA) problem. This mechanism for channel assignment on multi-radios, i.e., forming grid-loops via Minimum Spanning Tree (MST) and forming group channel, which provides a novel mechanism to the Multi-Radio Multi-

Channel Wireless Mesh Networks (MRMC-WMN) for assigning channels to different loops and the related maintenance and renewing.

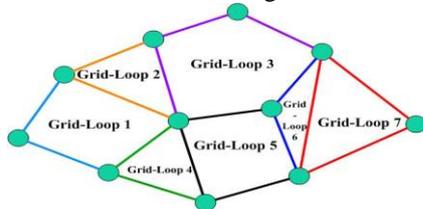


Fig3: Grid Loop Structure in MRMC

Channel Assignment in a MRMC-WMN environment recognizes efficient channel utilization and minimizes intervention but limited number of channels accessible to the network [2].

3.2, CACAO Algorithm

The distributed CACAO algorithm cultivates information collected by APs and clients on interference circumstances to diminish the local objective function by switching to a channel that has minimum predictable interference. At every time period, each AP individually elects the channel assignment to minimize the objective function locally. Obviously, the whole network does not need any synchronization.

Algorithm 1 CACAO Algorithm

CACAO (*api*)

1. Initialization - Initial Assignment
 $api.c \leftarrow rand(k)$
2. Optimization - Repeated for each AP
 2. a Gather Statistic ()
 2. b $ct = Compute\ Interfere ()$
 2. c Switch To (*ct*)

Gather Statistic () is a procedure that is used to gather statistics from clients. It returns the traffic information collected by clients. Then the AP performs computation and comparison to decide which channel it will use for the next time period by procedure Compute Interfere (). This routine computes the expected significance of interference level that the entire BSS will likely experience for the next time interval. The channel with the least obstructed traffic will be chosen. Finally the Switch To () routine allocates the AP and its associated clients the channel.

The limitations of CACAO algorithm randomly selected clients to switch to other channels for a small fixed period of time. AP only measures its operating channel for efficiency.

3.3, ILP Approach

In [4], the problems of channel assignment and AP placement are solved

Instantaneously by using Integer Linear Programming (ILP). The methodology considers not only radio coverage but also load balancing between APs because the authors argue that the number of active wireless clients connected to the Aps affects network performance. That is, traffic congestion at the APs degrades the network performance such as throughput. The basic idea is therefore to distribute clients to the Aps in a WLAN such that congestion at APs is minimized. Correspondingly, the throughput is maximized. A floor plan is assumed to consist of traffic demand points, each of which is given an expected traffic demand volume. A set of AP candidate locations is also given. If a signal from the AP to the demand point is above a certain threshold, an edge is drawn between a traffic demand point and a particular AP. Similarly, an edge is drawn between two APs, whenever they are within a co-channel interference distance defined as a transmission range at which, if assigned the same channel, these two APs can interfere to some extent with one another. The objective is to minimize the maximum channel utilization at each AP, while keeping a certain level of traffic demand satisfied at each demand point. Each demand point is assigned to exactly one AP. If at least one demand point is assigned to an AP that AP will be included in the solution set. If an edge exists between two APs, each AP will be assigned a different non-overlapping channel. As mentioned earlier, the goal is to distribute clients throughout the network such that the overall network throughput is maximized. This requires an accurate network layout containing the descriptions of demand points with estimated traffic, client distribution, and received signal levels at each demand location. In general, since such a network layout is very dynamic, new assignment of demand points to APs and Channels to APs are necessary. The new assignment however may cause a certain amount of disruption to client traffic. The authors propose another ILP that aims to minimize the amount of client traffic disruption due to the new assignment process, while maintaining the resulting channel utilization below that of the previous assignment.

3.4, Priority-Map Approach:

In [5], the channel assignment Problem is solved in tandem with the AP placement problem. A floor plan of interest is first divided into pixels. Each pixel is prioritized based on its traffic requirements, i.e., how much capacity the pixel may need and for how long. A highest-priority pixel is thus designated as one having highest demand of capacity and availability, and similarly lowest-priority pixel having lowest demand of capacity and Availability. Intermediate levels of priority are possible. With the priority map created; the strategy is to first come up with a set of possible AP locations which can provide adequate radio coverage to every pixel of the floor plan. Adequate coverage here means a minimum level of capacity and availability required by each pixel. To achieve this, a wave propagation prediction model such as a ray tracing technique can be used to predict the coverage area generated by each candidate AP. As can be expected, more than one set of possible AP locations may result from this process. For each set of possible AP locations, the next step is to eliminate those APs which create unacceptably large coverage overlap between their adjacent neighbors,

provided that adequate capacity and availability are still supplied to every pixel affected by this elimination. To quantify this overlap, the mean difference between the received powers of two adjacent APs is used. That is, if the difference in received power averaged over every pixel in the overlap area falls below a certain threshold, one of the APs should be eliminated. This step thus eliminates the possible interference created by the radio coverage overlap of adjacent APs. After the above elimination process, the channel assignment is now applied to each set of possible AP locations. The assignment starts in a greedy manner in which a no overlapping channel is assigned first to the AP that covers the area with highest priority.

TABLE 1
SUMMARY OF CHANNEL ASSIGNMENT
TECHNIQUES IN IEEE 802.11

Techniques	Nature		Deploy		Channel		References
	static	Adaptive	Uncoordinated	Centrally Managed	Overlap	Non Overlap	
GDCA Algorithm	*			*		*	[2]
CACAO Algorithm	*			*		*	[3]
ILP Approach	*			*		*	[4]
Priority-Map Approach	*			*		*	[5]

The challenge would be how to capture the network dynamics as much as possible while maintaining the complexity of implementation of channel assignment algorithm at a practical level. Furthermore, when WLANs are deployed in an uncoordinated fashion by different network administrators, the scalability of the implementation of channel assignment algorithms becomes even more important issue. In such scenarios, a channel assignment scheme of choice should be cooperative and Scalable enough to orchestrate channel switching across the entire network without creating significant interference to the neighbors. Being aware of the neighboring networks located in different administrative domains, the scheme should also be able to interact and exchange necessary information with its neighbors in order to allocate appropriate channels to the APs.

The research direction tends to shift toward adaptive channel assignment in uncoordinated environments, in which network dynamics is incorporated

IV. Conclusion And Open Research Issues

Channel assignment is one mechanism to improve the performance of WLANs. In this survey we have discussed several existing channel assignment schemes applicable to either centrally managed or uncoordinated environments. Several possible future research directions and open issues with regard to channel assignment in WLANs are outlined below:

into the problem formulation. The following system parameters need to be considered: client locations, building layouts and AP locations, time fluctuation of traffic demand of wireless clients at various locations, and application Quos requirements.

Continually monitoring the network dynamics, say on a daily basis, at a particular location may lead to a discovery of traffic pattern. Channel assignment can then be performed at a particular location during a particular period of time based on the prediction as well as the application requirements. The schemes discussed in this survey assume either Uplink or downlink traffic. To be more realistic, traffic in both directions should be considered. This is reasonable as peer-to-peer communications become more popular.

References

- [1] F. Cali, M. Conti, and E. Gregorian, "Dynamic tuning of the IEEE 802.11 Protocol to achieve a theoretical Throughput limit," *IEEE/ACM Trans. Networking*, vol. 8, no. 6, pp.785-799, Dec. 2000.
- [2] Yan Xia, Yingzhi Zeng, Zhenggu Gong, "A Distributed Channel Assignment Mechanism for Multi-radio Multi-Channel Wireless Mesh Networks" *JOURNAOF NETWORKS*, VOL. 6, NO. 2, FEBRUARY 2011. Xiaonan Yue, Chi-Fai Wong, S. -H.Gary Chan. "A Distributed Channel Assignment Algorithm for Uncoordinated WLANs" *IEEE CCNC 2010 proceedings*.
- [3] Y. Lee, K. Kim, and Y. Choi, "Optimization of AP Placement and Channel Assignment in Wireless LANs," in *Proc. 27th Annual IEEE Conf. Local Computer Networks*, pp. 831-836, Nov. 2002.
- [4] P. Wertz, M. Sauter, F. Landstorfer, G.Wolfle, and R. Hoppe, "Automatic Optimization Algorithms for the Planning of Wireless Local AreaNetworks," in *Proc. IEEE Vehicular Technology Conference*, vol. 4, pp. 3010-3014, Sep. 2004.
- [5] A. Mishra and V. Shrivastava, D. Agrawal, S. Banerjee, and S. Ganguly, "DistributedChannel Management in Uncoordinated Wireless Environments," in *Proc.International Conference on Mobile Computing and Networking*, pp. 170-181, 2006.
- [6] I.KatzelaandM.Naghshineh,"ChannelssignmentSchemesfor CellularMobileTelecommunicationSystems:AComprehensiv e Survey,"*IEEE Personal Commun.*, June 1996, pp. 10-31
- [7] S. Pack, J. Choi, T. Kwon, and Y. Choi,"Fast-Handoff SupportinIEEE802.11WirelessNetworks," *IEEE Commun. Surveys Tutorials*, vol.9,no.1,pp.2-12, 2007
- [8] P.Mahonen,J.Riihijarvi,and M. Petrova,"Automatic Channel Allocationfor Small Wireless Local Area Networks using Graph Colouring Algorithm Approach," in *Proc. IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 536-539, Sept. 2004
- [9] M. Gast, *802.11 Wireless Networks: TheDefinitive Guide* O'Reilly, Farnham, 2nd ed., 2005.

An Approach to Fractal image compression based on

Annapurna Misra, Nalini Singh, Boshisottva Dash

Abstract: The conception of digital image coding techniques is of great interest in various areas concerned with the storage or transmission of images. Fractal-based. Image coding or fractal image coding is a new method of image compression. In this method, similarities between different scales of the image are used for compression. Intuitively, fractals are sets that reveal details at every scale. These sets are in contrast to regular sets like lines, curves and planes that are typically studied in Euclidean geometry and become smooth when sufficiently magnified. In the early implementations of the theory, these similarities were found by human interaction and hence, the images were encoded by interactive computer programs. This resulted in codes for images that were extremely compact in size, but the decoded images had very low quality. In this paper, we propose an independent and novel approach to image coding, based on a fractal theory of iterated transformations.

Keywords: Euclidean geometry, Fractal, transmission, human interaction

I. INTRODUCTION

Fractal dimension has been used as a tool in different aspects of image compression algorithms. Fractal dimension was used in a fractal image coder for adjusting error threshold. Also fractal dimension was used for image segmentation. After segmentation, fractal dimension was used as a measure of the complexity of the segment to determine how the segments should be coded. To the human eye, many fractal curves and surfaces, look very similar to natural curves and surfaces, and for this reason they are extensively used in computer graphics. In model based coding, this similarity has the potential to be used for coding of natural images by modeling the processes that generate parts of these Images. In the context of image coding, fractal dimension was also used for selecting the optimal scale parameter in an edge detector. Experiments have shown that the fractal dimension of a curve or set is closely related to human perception of its roughness. Although fractal dimension alone is not enough for generating a visually good approximation of a set, it may be used as one of the parameters for its representation. It is a known fact that the human visual system's sensitivity to details -in any part of the image is dependent on the amount of activity in the background surrounding that part. Fractal dimension of image regions has been used as an objective measure of this activity.

II. LITERATURE REVIEW

1. Arnaud E.Jacquin in his paper "Image Coding based on a Fractal Theory of Iterated Contractive Image Transformations" described about the contractive image transformation and the coding techniques.
2. Geoffrey Davis of Department of Mathematics, Bradley Hall, Dartmouth College, Hanover, in his paper "Adaptive Self-quantization of Wavelet subtrees: A Wavelet based theory of Fractal image Compression" has described about the methods of Fractal image coding of wavelet transformed images.

III. FRACTAL IMAGE CODER

A. Encoding Steps

First ,the image is partitioned into a number of blocks of size $2B \times 2B$ pixels called domain blocks, which are used as building blocks, and a number of blocks of size $B \times B$ pixels referred to as range blocks. For each range block the aim is to find the domain block and transformation which is the best match of the original range block[1,2]. The transformation τ_i is defined as the transformation from domain block D_i to range block R_i and is written as the composition of two transformations where S_i and T_i are the so- called geometric and massic parts of τ_i , respectively as in equation(1).

$$\tau_i = T_i \circ S_i \quad (1)$$

A pool of domain blocks D is defined and consists of all image blocks of size $2B \times 2B$ which can be extracted from the original image, and a pool of massic transformations T , made of all block transformations T_i . The problem of finding the best match is then equal to the problem of finding the "best pair" $(D_i, T_i) \in D \times T$, which complies with the minimum distortion given in equation (2)

$$\delta_{12} (\mu\tau R_i, T_i \circ S_i (\mu\tau D_i)) \quad (2)$$

The transformation consists of two parts. The two parts are Geometric part and Massic part.

Geometric Part : A down scaling of the domain block in size from $2B \times 2B$ pixel to $B \times B$ pixels and a geometric displacement from the location of the domain block to the location of the range block.

Massic Part : Two transformations which manipulates the grey levels of the block and an isometric transformation which simply shuffles pixels within the block.

Contrast scaling by α :

$$(\sigma\mu)_{i,j} = \alpha\mu_{i,j} \quad (3)$$

Luminance shift by Δg :

$$(\tau\mu)_{i,j} = \mu_{i,j} + \Delta g \quad (4)$$

Isometries by I_i :The following transformations simply shuffle pixels within a range block, in a deterministic way- we call them isometries[3]. We list in the following a list of the eight canonical isometries of a square block as shown in the figure1.

- **Identity :**

$$(I_0\mu)_{i,j} = \mu_{i,j} \quad (5)$$

- **Orthogonal reflection about mid-vertical axis(j=(B-1)/2) of block:**

$$(I_1\mu)_{i,j} = \mu_{i,B-1-j} \quad (6)$$

- **Orthogonal reflection about the mid-horizontal axis (i=(B-1)/2) of block:**

$$(I_2\mu)_{i,j} = \mu_{B-1-i,j} \quad (7)$$

- **Orthogonal reflection about first diagonal (i=j) of block:**

$$(I_3\mu)_{i,j} = \mu_{j,i} \quad (8)$$

- **Orthogonal reflection about second diagonal (i + j = B - 1) of block :**

$$(I_4\mu)_{i,j} = \mu_{B-1-j,B-1-i} \quad (9)$$

- **Rotation around center of block, through + 90° :**

$$(I_5\mu)_{i,j} = \mu_{j,B-1-i} \quad (10)$$

- **Rotation around center of block, through + 180°:**

$$(I_6\mu)_{i,j} = \mu_{B-1-i,B-1-j} \quad (11)$$

- **Rotation around center of block, through - 90° :**

$$(I_7\mu)_{i,j} = \mu_{B-1-i,j} \quad (12)$$

In effect, massic transformations allow us to generate, from a single block, a whole family of geometrically related transformed blocks, which provides a pool in which matching blocks will be looked for during the encoding.

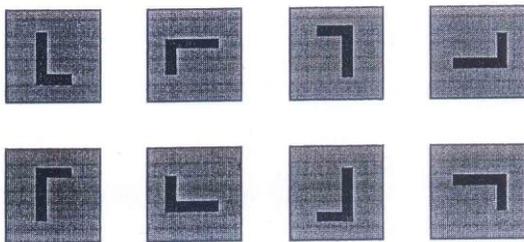


Figure 1. Different forms of Image using Isometries.

The total transformation from a domain block $\mu_{\tau D_i}$ to a range block $\mu_{\tau R_i}$ is given in equation (12)

$$T_i(S_j(\mu_{\tau D_i})) = I_i(\alpha_i(S_j(\mu_{\tau D_i})) + \Delta g_i) \quad (12)$$

With the δ_{l_2} metric we are able to calculate the contractivity of a transformation for example the contrast scaling on the block S (io, jo, B) as given in equation(13)

$$\delta_{l_2}(\sigma(\mu_{\tau s}), \sigma(v_{\tau s})) = \sum_{0 \leq i, j < B} (\alpha \mu_{i_0+i, j_0+j} - \alpha v_{i_0+i, j_0+j})^2 \quad (13)$$

$$= \alpha^2 \delta_{l_2}(\mu_{\tau s})$$

So the contractivity of this transformation is α^2 . Let μ is an image or a block and $\tilde{\mu}$ is the approximation of μ . A square block of size B x B pixels located at position (i_0, j_0) in μ is denoted by $s(i_0, j_0, B)$. To measure the distortion between the two blocks μ and $\tilde{\mu}$ of sizes B x B pixels, the following peak signal-to-noise ratio is given in equation (14) :

$$PSNR = 10 \log_{10} \left(\frac{255 * 255}{\delta_{l_2}(\mu, \tilde{\mu}) / B^2} \right) \quad (14)$$

The $\delta_{l_2}(\mu, \tilde{\mu})$ is the squared distortion between the two blocks is given in equation (15):

$$\delta_{l_2}(\mu, \tilde{\mu}) = \sum_{0 \leq i, j < B} (\mu_{ij} - \tilde{\mu}_{ij})^2 \quad (15)$$

B.Image Coding Algorithm Suppose that the image is segmented into blocks of size 4 X 4 pixels called Ranges. Each range block R must be approximated as $R \approx sD + O1$, where D is a 4 X 4 block from the shape codebook. Consider any domain block of size 8 X 8 in the image [4, 11]. Then shrink the block by pixel averaging to the desired size of 4 X 4 pixels. All such blocks are added to the shape codebook. For an image of size 512 X 512 pixels this process yields a huge codebook with 255025 blocks. In order to reduce the number of blocks to a more manageable size one may consider only those blocks that have their upper left corner pixel on a regular square grid with a spacing 1:1. For example, with $l=8$ we would obtain 4096 adjacent domain blocks, which is often used in practice.

The encoder has to solve the following problem. For each range block the best approximation $\approx sD + O1$ needs to be found. In fractal encoding the coefficients s and o are called scaling and offset. To obtain optimal s, o and D, a scan of all the codebooks_blocks D should be performed. For each codebook block D the best coefficients s and o need to be determined. However, for all but the smallest scalar codebooks for s and o this is computationally infeasible. It takes too long. Fortunately, there exists a shortcut. If we work with the Euclidean norm when making the selection of the best coefficients i.e., finding the best coefficient.

$$E(D, R) = \min_{i_0, j_0} \square R - (sD + oI) \square \quad (16)$$

Given the two blocks R and D with n pixel intensities, r_1, \dots, r_n and d_1, \dots, d_n we have to minimize the quantity as in equation (17)

$$\sum_{i=1}^n (s \cdot d_i + o - r_i)^2 \quad (17)$$

The best coefficients are given by equation (18)

$$s = \frac{n(\sum_{i=1}^n d_i r_i) - (\sum_{i=1}^n d_i)(\sum_{i=1}^n r_i)}{n(\sum_{i=1}^n d_i^2) - (\sum_{i=1}^n d_i)^2} \quad (18)$$

And

$$o = \frac{1}{n} \left(\sum_{i=1}^n r_i - s \sum_{i=1}^n d_i \right) \quad (19)$$

With s and o given the square error is given in equation (20)

$$E(D, R)^2 = \frac{1}{n} \left[\sum_{i=1}^n r_i^2 + s(s \sum_{i=1}^n d_i^2 - 2 \sum_{i=1}^n d_i r_i + 2o \sum_{i=1}^n d_i) + o(o n - 2 \sum_{i=1}^n r_i) \right] \quad (20)$$

This procedure yields two real numbers s and o. For the encoding we can only use the quantized values from the scalar codebooks [5, 12]. Usually, one employs uniform scalar quantization amounting to a rounding operation.

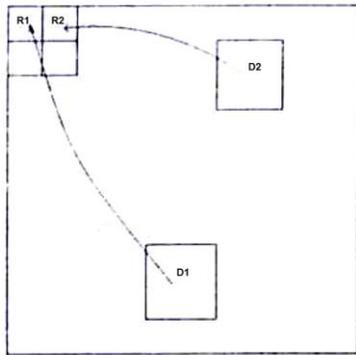


Figure 2. Example of a range and domain block

In summary the baseline encoder using a fixed block operates in the following steps.

1. *Image segmentation*: Segment the given image using a fixed block size, e.g., 4 X 4. The resulting blocks are called the ranges R_i .
2. *Domain pool and shape code book*: By stepping through the image with a step size of I pixels horizontally and vertically create a list of domain blocks from the image, which are twice the range size. By averaging four pixels each, shrink the domain blocks to match the size of the ranges. This produces the codebook of blocks D_i

3. The search for each range block R an optimal approximation $R \approx sD + oI$ is computed as follows. For each codebook block D_i compute an optimal approximation $R \approx s D_i + oI$ in three steps:

- Perform the least squares optimization using the given formulae, yielding a real coefficient s and an offset o.
- Quantize the coefficients using a uniform quantizer.
- Using the quantized coefficients s and o compute the error $E(R, D_i)$.

Among all the codebook blocks D_i find the block D_k with the minimum error [6, 7]. Output the code for the current range block consisting of indices for the quantized coefficients sand o and the index k identifying the optimal codebook block D_k . As already mentioned the output code of this baseline encoder is not a code with which the decoder can directly recover an approximation of the original. Instead we have a description of an operator.

Given the original image along with its partitioning in square ranges replace each range R by the corresponding blocks $sD + oI$ as specified by the code. The resulting image, called collage, is an approximation of the original. Thus the code is nothing but the description of an image operator. Given any image 90 one can carry out the operation iteratively to arrive at the reconstructed image. This is exactly what is done at the decoder. Usually the domain blocks are chosen to be twice the size as the corresponding range blocks. The contract activity condition for the image operator does not require a geometric contraction of domain blocks. Therefore the domain blocks can be of any size. It seems, however, that the error propagation at the decoder is generally worse when the geometric scaling factor is too small [8]. Therefore, shrinking 'the domains to half their original size is practical from the computational point of view and seems to produce the best looking results. It is common practice to enlarge the domain pool by including blocks obtained by rotation and by reflection.

IV. RESULTS AND DISCUSSIONS

In fractal coding usually a square block of size $2^r \times 2^r$ is approximated by another image block size of size $2^{r+1} \times 2^{r+1}$ under an affine mapping. Thus, one tries to find similar structures at two different scales. For a given 512 X 512 grey scale image, partitioned into non overlapping 16 X 16 blocks, a fractal code C is determined in the standard way, considering the domain pool non overlapping 32 X 32 blocks. C contains the information of 32.32 transformations [9,10]. In the decoding, C is used to compute the attractor A_1 of size 512 X 512. But C can also be iteratively applied to an arbitrary 256 X 256 image, partitioned into 8 X 8 blocks, gaining an attractor A_2 , or to an 128 X 128 image, partitioned into 4 X 4 blocks, giving an attractor A_3 , and so on. Thus, one ends up with a pyramid A_1, A_2, \dots, A_5 , describing different resolutions of the attractor.



Figure 3. Output of the coder showing results for different iterations(a,b,c,d,e,f,g) for iterations(1,2,3,5,7,10,20)

No. of iterations of Lena image	Encoding Time in sec	Decoding Time in sec	MSE	PSNR
1	62.891	1.437	521.6	69.1
2	62.078	2.86	289.6	71.7
3	62.922	4.297	159.0	74.3
5	62.719	7.109	107.9	75.9
7	62.672	9.953	107.9	75.9
10	63.328	14.265	107.9	75.9
20	63.156	28.516	107.9	75.9



Figure 4. Output of fractal coder of Lena image for one iteration

The result shows applying iterations on the image, for a less no. of iterations the MSE is more and for more iteration MSE is less. After certain iterations we get a constant MSE and PSNR value.

V. CONCLUSION AND FUTURE WORK

When the above theory is used for image compression, it is implemented in a discrete setting; however, the fractal code generated by encoding a digital image describes relationships, in the form of affine functions, between various segments of the image and is independent of the resolution of the image. In other words, the fractal code is a resolution independent representation of the image and theoretically represents a continuous image approximating the original image. A decoder may decode the code to generate a digital image at any resolution. The resolution of the decoded image may as well be higher than the resolution of the original image. The increase in resolution is sometimes referred to as the fractal zoom. The higher resolution obtained is not created by a simplistic technique such as repeating the pixels of the image, but more detail is actually generated in the decoded images. When an image is reconstructed at the same resolution as the original, in the decoding process the domain blocks are shrunk, which eliminates some of the details of the domain blocks. However, if the image is reconstructed at a higher resolution, in the shrinking of the domain block, the details of the domain block are only shrunk to generate the extra resolution in the range block. In fact, details of the domain blocks are used for missing details of the range blocks. The details in the domain block are also generated to some extent

from details of other domain blocks used for encoding each part. In other words, it is implicitly assumed that if the range block is similar to its corresponding domain block, then the details of the range block, which are beyond the resolution of the original image, are also similar to the details of the domain block, which are within the resolution of the encoded image. This assumption is a typical property of self similarity of fractal sets at different scales and the resolution independence is a property of the code generated by the fractal based methods. There are a number of ways in which the designed codec can be improved. Ongoing research around the world has suggested combination of more than one scheme for achieving fast, deep and less distortion compression. Among some of these schemes are a combination of wavelets and vector quantization, fractals with vector quantization and even wavelets with neural networks. The coder was designed for grey scale images only. The same algorithm could be used for color images too with modifications. Fractal coding is also being tried for video coding and has proved to be promising. Further work in this direction should consider efficient compression using zero trees, compression of color images, and deep video compression.

ACKNOWLEDGMENT

The authors thank to Dr. R. N. Pal, retired Professor IIT Kharagpur, India and his team for their help in conducting experiments with different face images.

REFERENCES

1. Fractals and the human visual system," in Proceedings of IEEE ICASSP-90, Albuquerque, NM, Apr.3 {6,1990, pp. 1957-1960}.
2. J. E. Hutchinson, Fractals and self-similarity," Indiana University Mathematics Journal, vol.30, pp. 713{747, Sept.-Oct.1981}.
3. M.F. Barnsley, Methods and apparatus for image compression by iterated function systems." United States Patent Number 4, 941, 193, 1990.
4. C.K. Lee and W.K. Lee, "Fast fractal image block coding based on local variances", IEEE Trans. Image Processing, Vol 7, No.6, pp 888-891, June 1998.
5. Y. Fisher, Ed., Fractal Image Compression: Theory and Applications;, New York: Springer- Verlag 1995.
6. A. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations", IEEE Trans. Image Processing, Vol.1 No.1, pp 18-30, Jan 1992.
7. G.M. Davis, "Self-quantization of wavelet subtrees: A wavelet-based theory of fractal image compression. "in Proc. Data Compression Conf. Snowbird. UT. Mar, 1995. pp. 232-241.
8. H. Krupnik. D. Malah and E. Kamin., "Fractal representation of images via the discrete wavelet transforms." In IEEE 18th Conv. Electrical Engineering in Israel. Tel-Aviv. Israel, Mar 1995.
9. A. van de Walle. "Merging fractal image compression and wavelet transform methods." In Fractal Image Coding and Analysis: A NATO ASI Series Book. Y. Fisher. Ed. New York: Springer-Verlag. 1996.
10. J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients." IEEE Trans. Signal Processing. Vol.41. pp. 3445-3462. Dec., 1993.
11. Y. Fisher, B. Jacobs, and R. Boss, "Fractal image compression using iterated transforms." In Image and Text Compression. J. Storer. Ed. Boston, MA: Kluwer, 1992. pp. 35-61.
12. Fractals and the human visual system," in Proceedings of IEEE ICASSP-90, Albuquerque, NM, Apr.3 {6, 1990, pp. 1957-1960}.

Optimization of cryogenic treatment parameters to maximise the tool wear of HSS tools by Taguchi method

M. C. Kumar¹, Dr. P. VijayaKumar², Dr. B. Narayan³

1. Dept. of Mech. Engg., City Engineering College, Bangalore

2. Dept of Mech. Engg., UVCE, KR Circle, Bangalore

3. Dept of Mech. Engg, Dr. Ambedkar Institute of Technology, Bangalore

Abstract: This work describes the optimization of cryogenic treatment (CT) parameters such as the cooling rate, the soaking temperature, the soaking time, the tempering temperature and the tempering time to maximise the tool wear of high speed steel tools in dry turning of AISI 316 grade austenitic stainless steels by Taguchi method. For this purpose, four iteration of Taguchi design (L16) have been used to arrive at the optimum CT parameters. The various levels of cryogenic treatment parameters have their own influence upon the flank wear resistance of the HSS tools. In this study, the Taguchi method has been used to optimize the process parameters for cryogenic treatment of commercial HSS tools. For optimization, four iterations of Taguchi design have been used to arrive at the optimum CT parameters. The experimental results demonstrate that the soaking temperature was the major parameter influencing the flank wear rate for all samples, but, other parameters were found to have negligible effect on the flank wear rate.

I. Introduction

High speed steel (HSS) is inexpensive, cutting tools can be shaped easily and it has excellent fracture toughness along with fatigue resistance. Due to its limited wear resistance it can be used at lower cutting velocities of 30-50 m/min [1]. Consequently, change in the chemical properties of HSS tool material to resist wear from the adverse machining environment. Cryogenic treatment (CT) shows promise as a tool material treatment for increasing tool life[2-5]. In this regards cutting tool industries have taken lot of interest on CT due to its effects on dimensional stability of the material, improv wear resistance and hardness of the materials. Many researchers [6-8] gave reasons for this improvement in properties because of complete transformation of retained austenite into martensite and the precipitation of fine carbide into the martensitic matrix.

Although CT is governed by many operational factors like cooling rate, soaking temperature, soaking time, heating rate, tempering temperature, tempering time etc, Various researchers [9-10] have used different levels of the CT parameters in their studies and have claimed different percentages of improvements in the mechanical properties of steel components [11] (Barron, 1982; Mohan Lal et al., 2001). Each CT parameter influences the quality of the tools for example the soaking temperature depends upon the martensite finish temperature of the material considered and soaking period depends upon the time required for the carbides to precipitate[12]. But no study was done to optimise the CT parameters for maximising tool wear resistance yet, because visualization of impact of various factors in interacting environment is not possible. On the other hand, an inexpensive and easy-to-operate experimental strategy based on Taguchi design has been adopted in this study to investigate the effect of various parameters and their interactions, because, in actual practice the resultant erosion rate is the combined effect of more than one interacting variable. This procedure has been successfully applied for parametric appraisal in different process optimization problems [13-14]. Further, the analysis of variance (ANOVA) is done to identify the most significant control factors and their interactions. The present work was done to study the effect and optimise the CT parameters on tool wear behaviour of HSS tools in dry turning of AISI 316 grade austenitic stainless steels by Taguchi method.

II. Design of Experiment

The Design of experiments was assembled for the CT with the objective of achieving tool wear resistance considered as the response. Table 1 shows five factors and four levels used in the experiment. If four levels were assigned to each of these factors and a factorial experimental design was employed using each of these values, number of permutations would be 625. The fractional factorial design reduced the number of experiments to sixteen. The cooling rate (A), the soaking temperature (B), the soaking time (C), the tempering temperature (D) and the tempering time (E) were assigned to the 1st, 2nd, 3rd, 4th and 5th column of L16 array respectively. The orthogonal array of L16 type was used and is represented in Table 2. This design requires sixteen experiments with five parameters at four levels of each and the interactions were neglected. The S/N ratios were computed for wear rate in each of the 16th trial condition and their values are given in Table 2.

Table 1. Selected fermentation factors and their assigned levels

\Symbol	Factors	Unit	Levels			
			1	2	3	4
A	Cooling rate	°C/min	1	2	3	4
B	Soaking temperature	°C	-80	-115	-150	-185
C	Soaking time	h	06	12	18	24
D	Tempering temperature	°C	125	150	175	200
E	Tempering time	h	1	2	3	4

Table 1 shows five factors and four levels used in the experiment. If four levels were assigned to each of these factors and a factorial experimental design was employed using each of these values, number of permutations would be 625. The fractional factorial design reduced the number of experiments to sixteen. The cooling rate (A), soaking temperature (B), Soaking time (C), Tempering temperature (D) and Tempering time (E) were assigned to the 1st, 2nd, 3rd, 4th and 5th column of L16 array respectively. The orthogonal array of L16 type was used and is represented in Table 2. This design requires sixteen experiments with five parameters at each of these four levels. The interactions between main factors were neglected. The S/N ratios were computed for wear rate in each of the 16 trial conditions and their values are as given in Table. 2.

Table 2. Experimental data and sample statistics

Ex. No.	Main factors					Observed response (Wear rate x10-5 mm ³ / m)			Average wear rate	Standard Deviation	S/N ratio (db)
	A	B	C	D	E	1	2	3			
1	1	1	1	1	1	4.4	4.553	3.451	4.135	0.597	-16.809
2	1	2	2	2	2	1.657	2.187	0.755	1.533	0.724	-6.516
3	1	3	3	3	3	3.45	3.876	3.132	3.486	0.373	-19.405
4	1	4	4	4	4	9.781	10.392	9.336	9.836	0.530	-25.368
5	2	1	2	3	4	3.37	3.81	2.955	3.378	0.428	-17.954
6	2	2	1	4	3	4.343	4.679	3.801	4.274	0.443	-19.689
7	2	3	4	1	2	4.173	4.34	3.382	3.965	0.512	-17.784
8	2	4	3	2	1	5.402	5.965	5.084	5.484	0.446	-21.792
9	3	1	3	4	2	4.379	4.535	4.257	4.390	0.139	-29.968
10	3	2	4	3	1	2.718	2.849	1.937	2.501	0.493	-14.105
11	3	3	1	2	4	4.466	5.097	3.634	4.399	0.734	-15.556
12	3	4	2	1	3	7.059	7.613	6.997	7.223	0.339	-26.566
13	4	1	4	2	3	5.102	5.928	4.77	5.267	0.596	-18.921
14	4	2	3	1	4	4.338	4.866	3.803	4.336	0.532	-18.231
15	4	3	2	4	1	5.435	5.784	4.534	5.251	0.645	-18.214
16	4	4	1	3	2	8.413	8.925	8.24	8.526	0.356	-27.581

III. Experimental studies

In this study, the HSS tools were cryogenically treated under dry conditions where the tools being treated were not exposed to the liquid nitrogen to eliminate the risk and damage of thermal shock. The procedure used for the treatment in this study is outlined in the following steps and is shown in Fig.1. Tools were placed in a container and the temperature was brought to -196°C in intervals by computerized controlled different rates. The temperature was held for different durations before the process was reversed. The tools were slowly brought to room temperature allowing the material to stabilize. Then, the tools were subjected to tempering cycles to relieve the stresses induced by cryogenic treatment. This was accomplished by increasing the temperature to various temperatures and then slowly reducing the temperature back to room temperature.

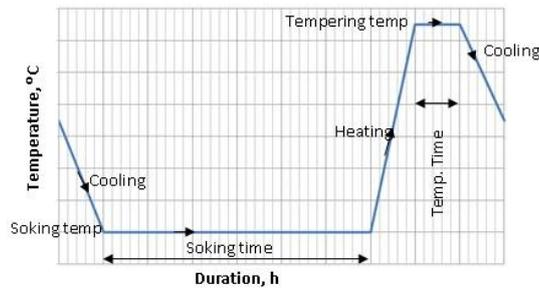


Fig. 1 Standard procedure for cryogenic and tempering treatment for HSS tools

Since the cutting mechanics involved in turning operation is relatively simpler, a difficult-to-machine material like stainless steel was considered for the turning tests were carried out on a HMT Heavy Lathe. For the experiment, cutting velocity, feed rate and depth of cut are 100 m/min, 0.2 mm/rev and 1 mm respectively. The duration of machining for each trial was only 60 s. After 60 s machining, the condition of the cutting tools was studied using optical microscopy. Once flank wear reached 0.3 mm, the tool life was considered to be over.

IV. Results and discussion

Optical micrograph (OM) was carried for both cryogenically treated and untreated HSS samples to study the microstructural changes. Results of the SEM analysis are shown in Fig. 2 (a) and (b) for untreated HSS tools and cryogenically treated tools respectively. The results showed the presence of the fine precipitated carbide particles in case of cryogenically treated samples which verify that the refinement of carbides takes place after the cryogenic treatment.

3.2 Taguchi's analysis of wear rate

Sixteen different set of experiments were performed using the design parameter combinations in the specified orthogonal array table. Three specimens were fabricated for each of the parameter combinations. The completed response table for these data appears in Table 2.

In order to estimate the effect of factor A (wt. % graphite) on the average value of response variable, three observed responses at level 1 of factor A, were summed together. Then the sum was divided by 4 to obtain the average response at level 1 of factor A. The average responses at level 2, 3 and 4 were obtained in the similar manner. The estimated effects are presented graphically in Fig. 3.

The estimated individual factors effect on average flank wear of HSS tools is shown in Fig. 3. The range of average responses over the four levels of each experimental factor, is: for cooling rate (A), 2 °C / min, for soaking temperature (B), -175 °C, for soaking time (C), 12 h, tempering temperature (D), 150 °C and for tempering time (E), 1 h. In particular, factor A (2), factor B (3), factor C(2), factor D (2) and factor E(1) for minimum flank wear combinations were established,

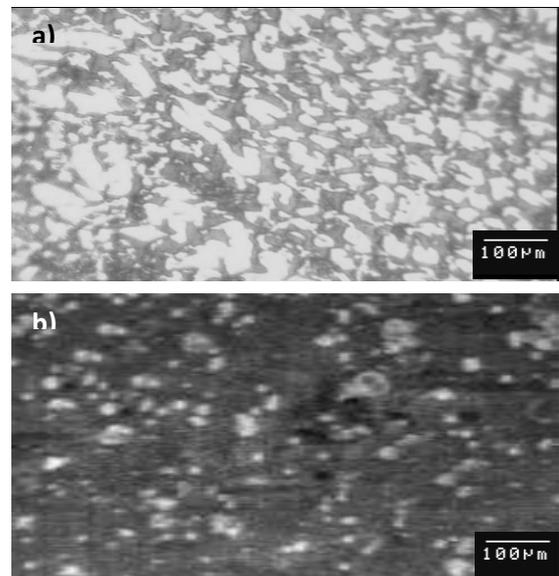


Fig. 2. Microstructure of a) untreated and b) cryogenically treated HSS tools

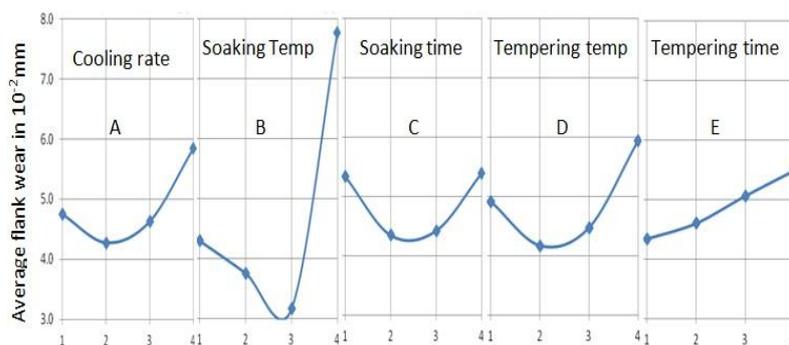


Fig. 3. Main effect and interaction effect plots for HSS samples

From Fig. 3, it can be seen that second level of factor (A) give the lowest average (i.e A2, which average of responses at cooling rate is 2 °C/min). The lowest average for factor B is at the third level (i.e B3, which is -175 °C), the lowest average soaking time (i.e C2 is 12 hour), the lowest average tempering temperature (i.e D2 is 150 °C) and the lowest average tempering time (i.e E1 is 1 hour). These predicted parameters are not used in the cryogenic treatment which is indicated in Table 2. Hence experiments were conducted at the predicted parameters, i.e cooling rate 2 °C/min (A2), soaking temperature -175 °C (B3), soaking time 12 hours (C2), tempering temperature 150 °C (D2) and tempering time 1 hour (E1), tested the specimen by flank wear test. The resulting wear rate was 1.323 mm, which is lower than the average flank wear in Table 3.

Table 3 Summary ANOVA Table for wear rate of PA66/graphite composites

Factors	DOF	Sum of squares	Mean squares	Fcal	Ftab	Alpha	% of Confidence	% of contribution
A	3	5.51	1.84	0.3886	1.52	0.7629	7.77	23.71
B	3	51.30	17.10	3.6189		0.0381	72.38	96.19
C	3	3.84	1.28	0.2711		0.8452	5.42	15.48
D	3	7.16	2.39	0.5049		0.6847	10.10	31.53
E	3	3.07	1.02	0.2165		0.8834	4.33	11.66
SW	15	70.88	4.73					

Examination of the calculated Fisher's values (F) for all control factors also shows a very high influence of factor B and low influence of factor E on the flank wear of the HSS samples as shown in Table 3. The F value was calculated for each design parameters. The optimum test conditions were estimated from the significant factors. The computed value of Fcal (3.6189) was more than that from the statistical Ftab (1.52). The Ftab equals to 1.52 at 99.5% confidence level. Thus, based on the level of confidence (99.5%), factors of only B (96.19 %) were significant while the factors A (23.71%), C (15.48 %), D (31.53%) and E (11.6 %) were less significant. The last column of the above table indicates the percentage of contribution (%P) of each factor, thus exhibiting the level of influence on the quality characteristic. The table shows that cooling rate (A), soaking temperature (B), for soaking time (C), tempering temperature (D) and tempering time (E), have percentage of contributions of 7.77 %, 72.38%, 5.42%, 10.10 % and 4.33 % on the flank wear respectively.

V. Conclusions

The Taguchi and ANOVA methods were applied to investigate the effects of cooling rate, soaking temperature, soaking time, tempering temperature and tempering time. The conclusions drawn from this work are;

- ♣ Cryogenic soaking temperature is the most significant factor and the maximum percentage contribution of soaking temperature on the flank wear of HSS tools was 72.38 %. The best soaking temperature in the possible range is -175 °C
- ♣ The second significant factor is tempering temperature and its contribution is 10.10 % for the improvement of flank wear resistance. The best level for this factor 150 °C.
- ♣ The other factors such as cooling rate (7.77 %), soaking time (5.42 %) and tempering duration (4.33%) contribution to wear resistance.
- ♣ The optimum test condition, at which the lowest rate is obtained, has been determined to be A2B3C2D2E1 levels

References

1. GogteCL, KumarMIyer, ParetkarRK and PeshweDR, DeepSubzeroProcessingofMetalsandAlloys: EvolutionofMicrostructu reof AISI T42ToolSteel, *Materials&ManufacturingProcesses*, vol. 24(2009), pp. 718-722.
2. GogteCL, KumarMIyer, ParetkarRK and PeshweDR, DeepSubzeroProcessingofMetalsand Alloys: Evolution of Microstructure of AISI T42 Tool Steel, *Materials & Manufacturing Processes*, vol. 24, (2009) pp. 718-722.
3. Yong, A.Y.L., Seah, K.H.W., Rahman, M., Performance evaluation of cryogenically treated tungsten carbide tools in turning, *Int. J. Mach. Tools Manuf.*, vol. 46(2004), pp. 2051–2056.
4. Yong, A.Y.L., Seah, K.H.W., Rahman, M., Performance of cryogenically treated tungsten carbide tools in milling operations, *Int. J. Adv. Manuf. Technol.*, vol. 32, (2007) pp. 638–643.
5. Ramji.B.R ,Narasimha Murthy.H.N ,Krishna.M, Performance analysis of cryogenically treated carbide drills in drilling white cast iron *International Journal Of Applied Engineering Research*, vol. 1, (2010) pp. 553-560
6. J.D. Darwina., D. Mohan Lalb, I, G. Nagarajanb, I Optimization of cryogenic treatment to maximize the wear resistance of 18% Cr martensitic stainless steel by Taguchi method *journal of materials processing technology*, vol.195(2008) pp. 241–247
7. Meng, F., Tagashira, K., Sohma, H., Wear resistance and microstructure of cryogenic treated Fe–1.4Cr–1C bearing steel. *Scripta Metall. Mater...*, vol.31(2005), pp.865–868.
8. Mohan Lal, D., Renganarayanan, S., Kalanidhi, A., Cryogenic treatment to augment wear resistance of tool and die steels. *Cryogenics* vol.41(2001), pp.149–155.
9. Seah, K.H.W, Rahman, M., Yong, K.H., Performance evaluation of cryogenically treated tungsten carbide cutting tool inserts, *Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf.*, vol. 217(2003), pp. 29-43.

10. Selvaraj D.P., Chandramohan P. (2010) , Influence of Cutting Speed, Feed Rate and Bulk Texture on the Surface Finish of Nitrogen Alloyed Duplex Stainless Steels during Dry Turning. Engineering, vol.2, (2010) pp. 453-460
11. Reddy, T.V.S., Sornakumar, T., Reddy, M.V., Venkataram, R., Senthilkumar, A. Turning studies of deep cryogenic treated p-40 tungsten carbide cutting tool inserts. Technical Communication .J. Mach Sci tech. vol. 13(2009), pp. 269-281
12. Paul S, Dhar NR, Chattopadhyay AB., 2001, Beneficial effects of cryogenic cooling over dry and wet machining on tool wear and surface finish in turning AISI 1060 steel, J Mater Process Tech, vol.116 (2001), pp 44–48.
13. S. Thamizhmanii, S. Saparudin, S. Hasan, Analyses of surfaces roughness by turning process using Taguchi method, Journal of Achievements in Materials and Manufacturing Engineering 20 (2007) 503-506.
14. D.H. Wu, M.S. Chang, Use of Taguchi methods to develop a robust design for the magnesium alloy die casting process, Materials Science and Engineering 379 (2004) 366-371.



M.C. Kumar, M.E is an Associate professor in the department of Mechanical engineering, City Engineering College, Bangalore, he has published 10 national and international papers.



Dr. P. Vijaya Kumar, M.E, Ph.d, is a Prof. and Dean of the faculty for Engineering studies at Department of Mechanical Engineering University Visvesvaraya College of Engineering, Bangalore University, Bangalore, he has 25 years of teaching both UG & PG, and He has published about 30 papers in national and international journals.



Dr. B. Narayan. M.E, Ph.d, is a Prof. in the Department of Mechanical Engineering, Dr. Ambedkar Institute of Technology, Bangalore he has 27yrs of teaching both UG & PG, he has published about 17 papers in national and international journals.

Effect of Processing Parameters on Polypropylene Film Properties

Ikilem Gocek¹, Sabit Adanur²

¹(Department of Textile Engineering, Istanbul Technical University, Turkey)

²(Department of Polymer and Fiber Engineering, Auburn University, USA)

Abstract- Thin mono-layer polypropylene (PP) films were produced to study the effects of processing parameters. Main difference in the crystalline structure of the PP films is considered to be caused by the orientation of the crystal blocks. As the temperature increases, film thickness decreases. Screw motor speed (SMS) increase results in increase of film thickness. Increase of winding speed (WS) causes the films to be stretched, making them thinner. As the temperature increases, tear strength of the films increases in MD but decreases in CD. Tear strength increases in CD with increasing SMS. As WS increases, tear strength in MD decreases. As the temperature increases, tensile strength increases in MD. When SMS increases, molecular orientation increases in MD causing tensile strength of the films to increase. Tensile strength decreases with increasing SMS in CD. Increase in WS causes tensile strength of the films to increase in MD and decrease in CD.

Keywords - Burst strength, polypropylene cast film, tear strength, tensile strength, winding speed

I. INTRODUCTION

Polypropylene (PP) finds various new uses in commercial film applications according to its product properties such as thickness, resistance to tear, tensile and burst, transparency, clarity, gloss and haze. Since PP has high tensile strength and elongation at break, low permeability and good chemical resistance as well as low price, it has been used in various packaging and other specialty applications [1].

The characteristics of cast PP films show differences not only due to their chemical structures and the additives and resin modifiers added but also due to their processing parameters in extrusion process. Therefore, determining the effect of processing parameters and process conditions in cast film extrusion is important to obtain the exact properties of the films for their end uses.

Cast film technology is the simplest technology to produce polymeric films in which the molten polymer is extruded through a slot die, fed by a single-screw extruder, onto a chilled roll in order to be cooled. Then, the solidified film is taken up from the chilled roll by a nip roll (take up roll) and transferred to a winding unit after the edges are trimmed. A limited amount of orientation is obtained in the film by this process; the orientation can be affected by the ratio between the die thickness and film thickness and by the ratio between the extrusion speed and the take up speed. Since the orientation can be easily determined in this method, cast film technology is a low cost and easy-to-handle process.

Temperature, screw motor speed and winding speed are the processing parameters that have effect on cast films. The final molecular orientation in a product depends on the thermomechanical history (melt temperature, stress, strain) of the product during processes [2]. It was reported that lower die temperature reduces the mobility of the chains to be extended at high draw ratios at the die exit [3]. Increasing screw motor speed results in the shear rate increase. At a given melt temperature, the higher the shear rate the higher the shear stress, which results in more molecular orientation in a product [2].

The aim of this study is to determine the effects of processing parameters such as temperature, screw motor speed and winding speed on the physical and structural characteristics and mechanical performance of unmodified and pure PP cast films such as thickness, crystallinity, tear strength, tensile strength and burst strength.

II. EXPERIMENTAL

2.1. Materials and Equipments Used

Polypropylene pellets commercially named as “30 Melt Copolymer Natural” with a density of 0.91 g/cm³ and an MFI of 34 g/10 min were bought from Premier Plastic Resins Company. No modifiers and additives were added while producing the PP cast films.

Wayne single screw extrusion machine having a 15.24 cm cast film die was used to produce polypropylene (PP) cast films in combination with a chill roll and a nip (take up) roll. The die melt was quenched onto a chill roll. The film was taken up through a pair of nip rolls onto a winder.

For tensile testing, Instron 5565 universal testing machine was used, which has 5 kN capacity and 0.001-1000 mm/min speed range. Thickness measurements were done using Testing Machines Inc., (TMI) micrometer. Crystallinity tests were performed on all film samples using a TA Instruments DSC Q-2000 differential scanning calorimetry (DSC).

2.2. Manufacturing and Testing

Twelve PP films were produced by using cast film technology. The machine parameters such as temperature, screw speed and winding speed were changed during production to observe their effects on film final properties. The machine parameters that were changed during the manufacturing process can be seen in Table 1. The barrel temperatures were the same along the barrel from feed end to the die for each sample.

Table 1 Machine parameters changed during manufacturing process of the films

Sample	Temperature (°C)	Screw Motor Speed (RPM)	Winding Speed (m/s)
Film1	190	400	0.1
Film2	205	400	0.1
Film3	215	400	0.1
Film4	220	400	0.1
Film5	225	400	0.1
Film 6	220	300	0.1
Film7	220	500	0.1
Film8	220	600	0.1
Film 9	220	400	0.12
Film 10	220	400	0.13
Film 11	220	400	0.15
Film 12	220	400	0.16

Thickness, tensile, tear and burst strength tests were performed on the samples. The effect of temperature on film thickness, tear, tensile and burst strength was investigated on films 1,2,3,4 and 5; the screw motor speed effect was investigated on films 4, 6, 7 and 8 and the winding speed effect was investigated on films 9, 10, 11 and 12. The tests were performed both in machine direction (MD) and cross direction (CD).

For tensile tests, “ASTM D 882-02 Standard Test Method for Tensile Properties of Thin Plastic Sheeting” test method was used. The nominal width of the specimens was 25.4 mm. The grip separation (gauge length) was 50 mm; the test specimens were 50 mm longer than the grip separation. The rate of grip separation (cross head speed) was 500 mm/min. Five specimens were tested from each sample for the tensile tests [4].

For tear tests, “ASTM D 1938-06 Standard Test Method for Tear-Propagation Resistance (Trouser Tear) of Plastic Film and Thin Sheeting by a Single-Tear Method” was used. The nominal width of the specimens was 25.4 mm; the length of the test specimens was 80 mm. The grip separation (gauge length) was 50 mm. The rate of grip separation was 250 mm/min. Five specimens were tested from each sample for the tear tests [5].

For burst strength tests, “ASTM D 6797-02 Standard Test Method for Bursting Strength of Fabrics Constant-Rate-of-Extension (CRE) Ball Burst Test” was used. The specimens have the dimensions of 125 x 125 mm. The CRE machine was started with a speed of 305 mm/min. and the speed was kept constant till the specimens bursted. Five specimens were tested from each sample for the burst tests [6]. For thickness tests, 10 measurements were taken from each sample. For crystallinity tests, samples were heated from -50 °C to 350 °C at a temperature increase of 10 °C/min.

III. RESULTS AND DISCUSSION

3.1. DSC Analysis and Crystallization

DSC tests were performed on all of the films. Since the melting peaks of all samples did not change much, it was concluded that the change of the crystal thickness was negligible (Fig. 1). Therefore, the main difference in the crystalline structure of the PP films is

thought to be caused by the orientation of the crystal blocks [3].

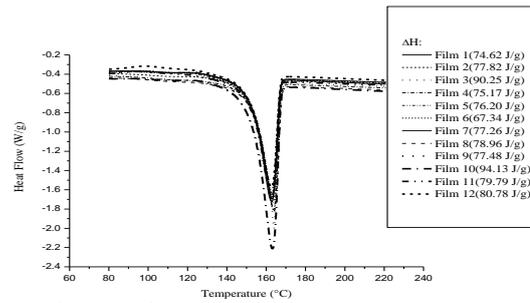


Figure 1 DSC of the PP films (heating rate: 10 °C/min)

Films 1, 2, 3, 4 and 5 were analyzed to examine the effect of temperature on crystallinity. Although the areas under the curves of DSC graphs and ΔH values (a parameter that can be used for defining crystallinity) fluctuate (Fig. 1), they show a slightly increasing trend. As the temperature is increased during PP film production process, the crystallinity of the films increases slightly as shown in Fig. 2.

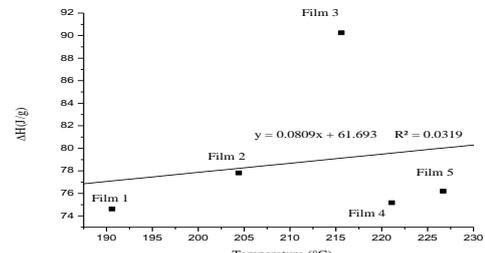


Figure 2 Effect of temperature on crystallinity

Films 4, 6, 7 and 8 were analyzed to examine the effect of screw motor speed on crystallinity (Fig. 3). As the screw motor speed increases, ΔH value increases.

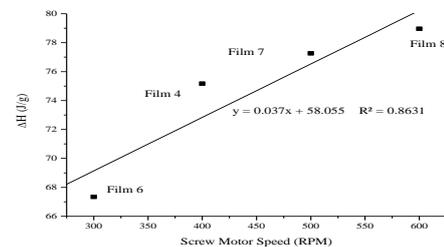


Figure 3 Effect of screw motor speed on crystallinity

Films 9, 10, 11 and 12 were analyzed to examine the effect of winding speed on crystallinity. Although the areas under the curves of DSC graphs and ΔH values fluctuate (Fig. 4), this is not enough to reach a conclusion about the winding speed and crystallinity relation.

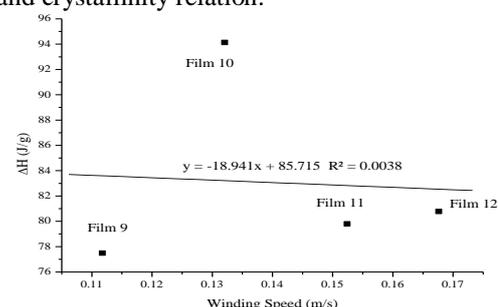


Figure 4 Effect of winding speed on crystallinity

3.2. Thickness Analysis

The results of thickness tests for all of the 12 PP film samples can be seen in Table 2.

Table 2 Thickness test results

Film #	1	2	3	4
Thickness (mm)	0.0355	0.0356	0.0321	0.0345
Film #	5	6	7	8
Thickness (mm)	0.0356	0.0224	0.0484	0.0522
Film #	9	10	11	12
Thickness (mm)	0.0332	0.0267	0.0221	0.0209

One of the extrusion machine parameters in film manufacturing process is temperature. Therefore, thicknesses of the films 1, 2, 3, 4 and 5 were measured to analyze the effect of temperature on thickness (Fig. 5). As the temperature increases, the thickness of the films decreases since the viscosity of the films decreases with increasing temperature.

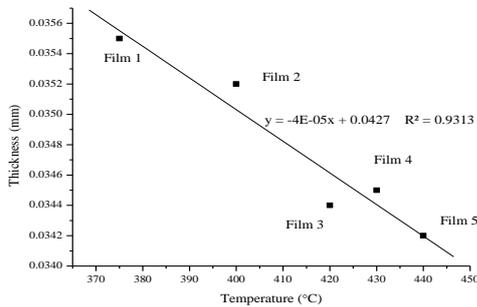


Figure 5 Effect of temperature on film thickness

Another extrusion machine parameter in film manufacturing process is the screw motor speed. Thicknesses of the films 4, 6, 7 and 8 were measured to analyze the effect of screw motor speed on thickness (Fig. 6). As the screw motor speed increases, the thickness of the films also increases since the screw transfers more polymer from the barrel to the die.

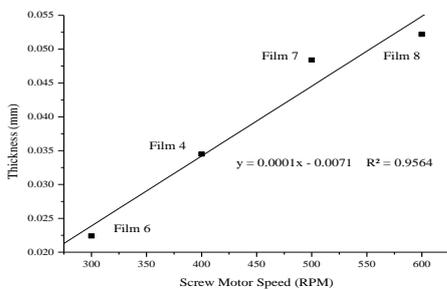


Figure 6 Effect of screw motor speed on film thickness

To analyze the effect of winding speed on thickness, thicknesses of the films 9, 10, 11 and 12 were measured (Fig. 7). As the winding speed increases, the thickness of the films drops since the increase in winding speed stretches the films and makes them thinner.

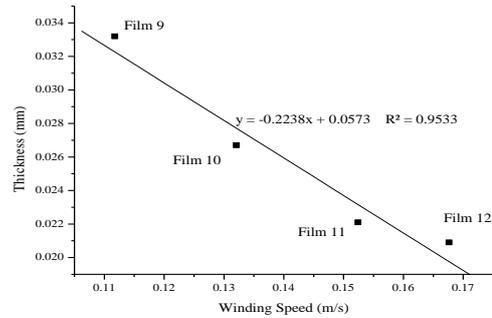


Figure 7 Effect of winding speed on film thickness

3.3. Tear Strength Analysis

Tear strength of the films 1, 2, 3, 4 and 5 were measured to determine the effect of temperature on tear strength (Fig. 8). As the temperature increases, tear strength of the films increases in MD. Since the polymer chains become more oriented in MD as the temperature decreases, tear strength becomes higher with increasing temperature. As the temperature increases, tear strength of the films decreases in CD since the polymer chains are less oriented in MD and less force is needed to tear the polymer chains in CD.

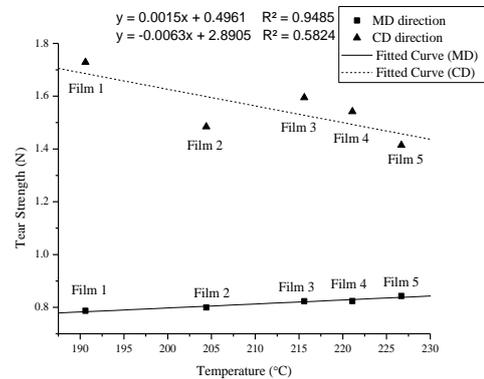


Figure 8 Effect of temperature on tear strength of the films

To analyze the effect of screw motor speed on tear strength, tear strengths of the films 4, 6, 7 and 8 were measured (Fig. 9). Although with increasing screw speed, the tear strength in MD is expected to decrease, the data do not show any specific trend. Therefore, no conclusion can be made from the data obtained. The shear rate increases with increasing screw motor speed. At a given melt temperature, higher shear rate gives higher shear stress, resulting in more molecular orientation in a product. The amount of molecular orientation is proportional to the magnitude of stress. Since the molecular orientation improves in MD, tear strength of the films increases in CD with increasing screw motor speed.

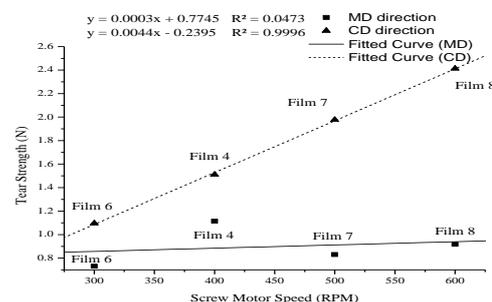


Figure 9 Effect of screw motor speed on tear strength of the films

To analyze the effect of winding speed on tear strength, tear strengths of the films 9, 10, 11 and 12 were measured (Fig. 10). As the winding speed increases, the tear strength of the films drops in MD. The orientation in MD increases due to high draw ratio. Therefore, tear strength in MD decreases with increasing winding speed. Although the tear strength of films is expected to increase in CD with increasing winding speed, the data show no specific trend.

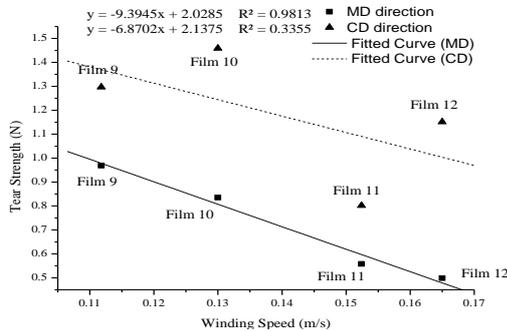


Figure 10 Effect of winding speed on tear strength of the films

3.4. Tensile Strength Analysis

Tensile strength of the films 1, 2, 3, 4 and 5 were measured to determine the effect of temperature on tensile strength (Fig. 11). At a given shear rate, lower melt temperature gives higher viscosity, resulting in higher shear rate. Lower melt temperature or higher shear rate gives rise to higher shear stress, resulting in more molecular orientation in MD. Therefore, as the molecular orientation increases due to lower temperatures, tensile strength increases. However, from Fig. 11, it can be seen that as the temperature increases, tensile strength increases both in MD and CD.

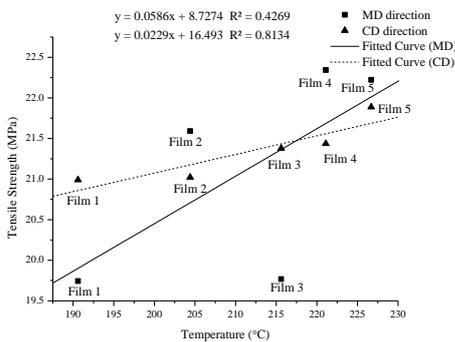


Figure 11 Effect of temperature on tensile strength of the films

To analyze the effect of screw motor speed on tensile strength, tensile strengths of the films 4, 6, 7 and 8 were measured (Fig. 12). The shear rate increases with increasing screw motor speed. At a given melt temperature, higher shear rate gives higher shear stress, resulting in more molecular orientation in a product. Since the molecular orientation increases in MD, the tensile strength of the films increases in MD. Tensile strength of the films decreases with increasing screw motor speed in CD because of the orientation in MD.

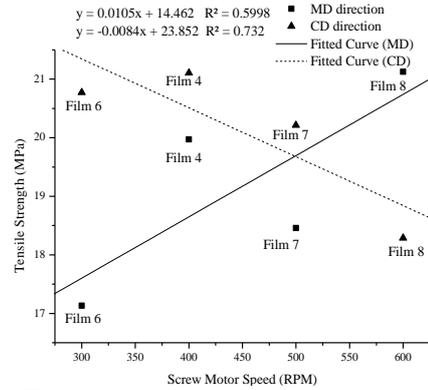


Figure 12 Effect of screw motor speed on tensile strength of the films

To analyze the effect of winding speed on tensile strength, tensile strengths of the films 9, 10, 11 and 12 were measured (Fig. 13). As the winding speed increases, the tensile strength of the films increases in MD since the orientation in MD improves due to the increasing draw ratio. The tensile strength of films decreases in CD with increasing winding speed due to the orientation increase in MD.

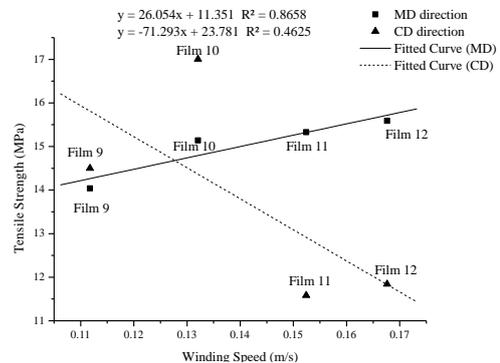


Figure 13 Effect of winding speed on tensile strength of the films

3.5. Burst Strength Analysis

During the burst strength tests, only Film 6 bursted. The tests were done by using Instron universal testing machine and there was not enough space in the machine's bursting strength equipment for highly stretchable films to extend. The graphs of the other 11 films did not show any decreasing trend at 45 mm extension. This means that the graphs obtained did not have their maximum loads since they have an increasing trend. Therefore, the burst strength values, stress at the maximum load, could not be obtained for the other 11 films.

As it is seen in Fig. 14, the graphs of all of the films show a similar trend. For the same extension value, 45 mm, Film 11 gives the lowest load value (not considering Film 6) while Film 8 gives the highest load value.

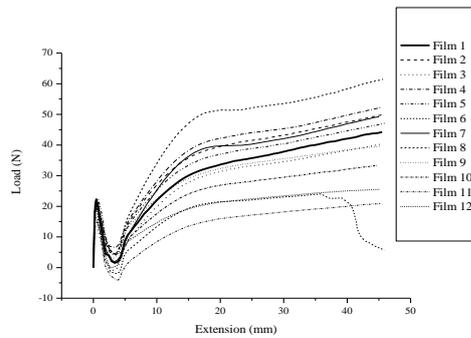


Figure 14 Burst strength of the films

The films exhibit early load peaks at the extension of nearly 0.4-0.8 mm range, but then their load values begin to decrease as the extension increases. The load values begin to increase from the extension range of 3.2-4.2 mm to the extension 45 mm.

IV. CONCLUSIONS

In this work, twelve different PP cast films were manufactured and analyzed to find the effects of machine parameters such as temperature, screw speed and winding speed on film properties such as crystallinity, thickness, tear strength, tensile strength and burst strength. These analyses were done in MD and CD directions for tear and tensile strength.

From the DSC analysis it was found that the position of the melting peak did not have a significant change. Because of this, it was assumed that the crystal thickness change was negligible. Therefore, the main difference in the crystalline structure of the PP films is considered to be caused by the orientation of the crystal blocks.

The thickness analysis was done by considering temperature, screw motor speed and winding speed. It was found that when the temperature increases, the thickness of the films decreases since the viscosity of the films decreases with increasing temperature. As the screw motor speed increases, the screw transfers more polymer from the barrel to the die which results in increase of film thickness. Increase of winding speed causes the films to be stretched, making them thinner.

Since the polymer chains become less oriented in MD as the temperature increases, tear strength becomes higher with increasing temperature. When the temperature increases, tear strength of the films decreases in CD since the polymer chains are less oriented in MD and less force is needed to tear the polymer chains in CD. Although it is expected that increasing screw speed decreases the tear strength in MD, tear strength of the films does not show any specific trend as the screw motor speed increases; as a result no conclusions can be made from the data obtained. Tear strength of the films increases in CD with increasing screw motor speed because the molecular orientation increases in MD. When the winding speed increases, the tear strength of the films decreases in MD because the

orientation in MD increases due to high draw ratio. Therefore, when winding speed increases, tear strength in MD decreases. Although the tear strength of films should increase in CD with increasing winding speed, the data show no specific trend.

When the molecular orientation increases because of lower temperatures, tensile strength should increase according to the literature. However in this work, as the temperature increased, tensile strength increased in MD. When screw motor speed increases, the shear rate also increases. At a given melt temperature, higher shear rate gives higher shear stress which causes higher molecular orientation in the product. Molecular orientation increase in MD causes the tensile strength of the films to increase. Tensile strength of the films decreases with increasing screw motor speed in CD due to the orientation in MD. Increase in winding speed causes the tensile strength of the films to increase in MD since the orientation in MD increases because of the increasing draw ratio. The tensile strength of films decreases in CD with increasing winding speed because of the orientation increase in MD.

Except Film 6, the films did not burst in burst strength tests. Therefore, burst stress at the maximum load could not be obtained for the other 11 films.

ACKNOWLEDGEMENTS

This research is supported by U.S. Department of Commerce (DOC-ITA-O8-TBD-E), which is appreciated. The authors would like to acknowledge Mr. William David Clark for his help during the production process of PP cast films.

REFERENCES

- [1] A. Christie, Polypropylene crystallization effects on film forming, www.optprocesssolutions.com, (09, 01, 2009).
- [2] C.I. Chung, *Extrusion of polymers theory and practice*, Hanser Verlag, (2000).
- [3] F. Sadeghi, A. Ajji, and P.J. Carreau, Analysis of microporous membranes obtained from polypropylene films by stretching, *Journal of Membrane Science*, 292, 2007, 62-71.
- [4] ASTM, Standard Test Method for Tensile Properties of Thin Plastic Sheeting, Designation D 882-02, vol. 8.01. American Society for Testing and Materials, Philadelphia, PA, (2002).
- [5] ASTM, Standard Test Method for Tear-Propagation Resistance (Trouser Tear) of Plastic Film and Thin Sheeting by a Single-Tear Method, Designation D 1938-06, vol. 8.01. American Society for Testing and Materials, Philadelphia, PA, (2006).
- [6] ASTM, Standard Test Method for Bursting Strength of Fabrics Constant-Rate-of-Extension (CRE) Ball Burst Test, Designation D 6797-02, vol. 7.01. American Society for Testing and Materials, Philadelphia, PA, (2002).

Financial Factor Affecting Maintenance Management In Safety And Health Practices

Hafizi Zakaria^{*}, Kadir Arifin, Shaharuddin Ahmad and Kadaruddin Aiyub
(School of Social, Development and Environmental Studies, Faculty of Social Sciences and Humanities,
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia)

ABSTRACT: Financial is often being the barrier to maintenance management in organizing maintenance activities of a building. This condition leads to a few maintenance management to neglect an accomplished maintenance although it will affect the safety and health of the building occupants. The aim of this paper is to present the current scenario of maintenance management in safety and health practice based on financial factor. Location of the study only focused on nine from sixty one blocks of government offices in Putrajaya, Malaysia. There are two groups of respondents in this study; the government office employees with a total number of 562 persons and the maintenance management employees with a total number of 100 persons. The study presents findings using quantitative methods including descriptive and inference statistics using Statistical Package for Social Sciences (SPSS) software. Majority of both respondents conclude that the maintenance management had realized their role towards the need of budget for every maintenance plan in safety and health practices. Besides that, the maintenance management is ready to provide sufficient budget for that purpose.

Keywords: Maintenance management, financial, safety, health

I. INTRODUCTION

Financial in the maintenance or more synonymous with the term maintenance budget could assure an effective and systematic maintenance planning. Financial means for the realization of necessary maintenance work should always be available at the right time. Recently, maintenance budgets were often not determined systematically, and a number of different budgeting methods are in used [1]. According to [2], maintenance is often regarded as a necessary expense that belongs to the operating budget. The maintenance budget should be balanced with the annual work schedule that has been decided. [3] add, budgetary control has been defined as the process whereby appropriate budgets are calculated and agreed, and expenditure against them monitored before and after commitment, to ensure that once budgets are set they are neither under-spent nor exceeded other than to meet variations in the performance criteria or market conditions on which they were based. Therefore, the appropriate budget should be prepared and designed based on the expected cost of maintenance work programmed in accordance with the requirements and the legal system or equipment. This is because the aspect of maintenance includes the maintenance of short term and long term.

Current building maintenance strategies, whether based on planned or unplanned maintenance, are most likely to be budget driven [4]. Rather than setting budgets to address needs, i.e. "zero-based budgeting" where operation and maintenance resources are estimated from scratch for

every fiscal year [5] [6]. [4] said the easiest way to cut maintenance costs is to stop doing maintenance and this approach is simple, but the long-term result are usually very costly. Most accountants still regard maintenance as a necessary evil that costs what it costs [7]. Thus, the investment (cost) in the maintenance gives the retention of assets value to help reduce the problem in terms of safety and health. [7] conclude that periodic maintenance was first prescribed to improve safety rather than to increase availability or reduce costs.

In Malaysia, the development plan allocation for repair and maintenance works increased from RM296 million during the Eighth Malaysian Plan to RM1,079 million in the Ninth Malaysian Plan and hence it is expected that repair and maintenance work will become more important in the future [8]. Maintenance and reliability of equipment are important factor that can considerably influence an organization's ability to compete effectively [9] [10]. [11] added, building maintenance is the combination of technical and administrative actions to ensure the items and elements of a building in an acceptable standard to perform its required function.

This paper aims to present the current scenario of maintenance management in safety and health practice based on financial factor. Among the key elements of the maintenance budget that effect the safety and health are adequate financial provisions for preventive maintenance and corrective maintenance repair including spare parts cost.

II. METHODOLOGY

This research focused on the office complex in Malaysia Federal Government Administrative Centre, Putrajaya. For the record, there are twenty one government office complex in Putrajaya with sixty one blocks of government offices in operation at present. It covers four main areas of Putrajaya Precinct 1, Precinct 2, Precinct 3 and Precinct 4. Nine blocks from sixty one government office blocks have been selected as the study site based on a few factor such as the group's office in each district or precinct, design, size or age of the area and office. According to [12] there are numerous factor that influence maintenance cost indicators of buildings to varying degrees including building characteristics, political and other factor.

2.1 POPULATION AND SAMPLE

Based on the research location selected, population and sample was divided into two groups. The population and sample for the first group is the government office employees of three categories; Top Management, Management and Profesional and Support (Table 1). This group is the user or recipient of services that will evaluate the overall performance of the financial factor of office maintenance carried out by maintenance management

organization. Thus, data and information from these groups are very useful in this study.

A total of eight top management representing 10 percent population of nine government offices involved in questionnaire survey. Meanwhile, 10 percent of the population from each government offices involved in questionnaire survey are represented by employees of management and professional and support.

The second group is the maintenance management employees. This group is the service providers who are able to determine the overall performance of maintenance work. A representative of this group consists of senior staff, middle staff and subordinates who have their respective functions in managing the maintenance office. Table 2 shows brief description about the functions of each category in the first group.

There are a total of fifteen maintenance management organizations in the nine research locations. Number of maintenance management organization in the research location is based on the division of the scope and responsibilities as determined by the maintenance management organization that appoints by the Ministry of Works Malaysia. There are three division in the scope and responsibility for maintenance management organization in the government offices, as illustrated in Table 3.

The scope and responsibilities of the first "comprehensive maintenance management services operation and maintenance of mechanical engineering and electrical engineering" involving a total of six maintenance management organization. While the scope and

responsibilities of "comprehensive maintenance management services operation and maintenance of civil engineering work, housekeeping, landscaping and pest control" also involves six maintenance management organization. Finally, the third scope and responsibilities of "comprehensive maintenance management services operation and maintenance of mechanical engineering, electrical, civil, housekeeping, landscaping and pest control" involves only three maintenance management organization.

Table 4 shows the number of population and sample for maintenance management employees. All of the fifteen senior staff of nine government offices were involved in the questionnaire survey. Meanwhile, 50 percent of the population of all government offices involved in the questionnaire survey are represented by the middle staff. Finally, 10 percent of the population of all government offices involved in the questionnaire survey are represented by the subordinate staff. According to [13], survey research sample should be at least 10 percent of the population.

2.2 DATA COLLECTION

Research requires systematic method in data and information collection to obtain accurate and robust research. This method was conducted to obtain data through a questionnaire survey. There are two parts in the questionnaire survey. Part 1 concerned about respondent personal data which consist of gender, age, duration of work in the office and job category. Part 2 is an assessment of respondent toward financial factor in office maintenance management.

Table 1. Number of population and sample for government office employees

Government office block	Number of population			Number of sample		
	Top management	Management and professional	Support	Top management	Management and professional	Support
A	11	252	181	1	25	18
B	8	136	467	0	14	47
C	7	174	491	1	17	49
D	10	291	341	0	29	34
E	9	198	231	1	20	23
F	10	269	505	1	27	51
G	10	287	472	0	29	47
H	8	97	361	1	10	36
I	8	274	510	3	27	51
<i>Total</i>	<i>81</i>	<i>1,978</i>	<i>3,559</i>	<i>8</i>	<i>198</i>	<i>356</i>

Table 2. Description of the main functions of maintenance management employees

Category	Designation	Function
Senior staff	Manager and Assistant Manager	Formulate, plan and manage activities related to maintenance management including preventive maintenance, corrective maintenance, cleaning, upgrading work etc.
Middle staff	Engineer, Executive, Technical Assistant and Supervisor	Monitor and implement activities related to maintenance management including preventive maintenance, corrective maintenance, cleaning, upgrading work etc.
Subordinates	Technician, Handyman, General Worker, Clerk	Implement activities related to maintenance management including preventive maintenance, corrective maintenance, cleaning, upgrading work etc.

Table 3. Scope and responsibility of maintenance management organization

Scope and responsibility		No.of maintenance management organization
1st	- Comprehensive maintenance management services operation and maintenance of mechanical engineering and electrical engineering	6
2nd	- Comprehensive maintenance management services operation and maintenance of civil engineering, housekeeping, landscaping and pest control	6
3rd	- Comprehensive maintenance management services operation and maintenance of mechanical engineering, electrical, civil engineering, housekeeping, landscaping and pest control	3
<i>Total of maintenance management organization</i>		<i>15</i>

Table 4. Number of population and samples for maintenance management employees

Government office block	Number of population			Number of sample		
	Senior staff	Middle staff	Subordinates	Senior staff	Middle staff	Subordinates
A	1	8	39	1	4	4
B	2	8	59	2	4	6
C	2	8	40	2	4	4
D	2	8	58	2	4	6
E	2	8	64	2	4	6
F	1	8	47	1	4	5
G	2	8	80	2	4	8
H	1	8	42	1	4	4
I	2	8	57	2	4	6
<i>Total</i>	<i>15</i>	<i>72</i>	<i>486</i>	<i>15</i>	<i>36</i>	<i>49</i>

2.3 DATA ANALYSIS

Analysis of survey data is done using the program Statistical Package for the Social Science (SPSS) version 16.0 to get the best results of the research. According to [14], SPSS is a computer software package that is very useful for data management and analysis. There are two types of analysis used in the study; descriptive analysis and inferential analysis. Descriptive analysis was used in data analysis in order to reveal and describe the phenomenon of the variables found in the survey questionnaire. That method aims to explore an area which has not been or less explored, with the aim to identify a thing to explain a phenomenon that is taking place. Descriptive analysis used in this study are described in terms of percentage (percent). The inference analysis used in this study is Mann-Whitney U test. The aim is to see the evaluation difference between two respondents for all items of financial factor in maintenance management.

III. RESULTS AND DISCUSSION

Table 5 shows the results of a questionnaire survey analysis Part 1. The research included a total number of 562 respondents representing government office employees and 100 respondents representing maintenance management employees in government offices, Putrajaya. Majority of the respondents or 358 government office employees are female which represent 63.7 percent, while the rest of them or 204 persons which represent 36.3 percent are male. For maintenance management employees, majority of them are male respondents or 82 persons which represent 82.0 percent and the rest of them are female or only 18 persons which represent 18.0.

The study also found that the largest percentage of respondents' age, which represents government offices employees are between 21 and 30 years, which reached 60.1 percent, and 20.3 percent of the respondents, age are between 31 and 40, followed by 10.0 percent of respondents, age are 51 years or more. Meanwhile, 9.1 percent of the respondents, age are between 41-50 years and only 0.5 percent of the respondents, age between 18-20 years or less. Next, the largest percentage of respondents age, which represents the maintenance management employees are between 21 and 30 years or 77.0 percent, while 16.0 percent are between 31 and 40, followed by 6.0 percent between 41 and 50 and only 1.0 percent age 51 years or more.

A total of about 181 persons or 32.2 percent of the government office employees have worked in the office for 1 to 2 years, followed by 20.6 percent (116 persons) of respondents have worked for 3 to 4 years, while 12.6 percent (71 persons) of respondents had worked for 6 to 11 months. Next, 11.8 percent (66 persons) of the government offices employees have been working for 5 to 6 years, 11.6 percent (65 persons) of respondents have worked for only 5 months or less and the rest only 11.2 percent (63 persons) of respondents have worked in the office for the longest period of 7 years or more. Most of the respondents' duration of employment in government offices, Putrajaya are due to the duration of the office development completion in stages since 1999. This description also takes into account that there were no migrations of the government office employees from the original government agency.

Table 5. Demographic profile of respondents

Description	Government office employees (%)	Maintenance management employees (%)
<i>Gender</i>		
Male	36.3	82.0
Female	63.7	18.0
<i>Age</i>		
18 - 20 years	0.5	-
21 - 30 years	60.1	77.0
31 - 40 years	20.3	16.0
41 - 50 years	9.1	6.0
51 years or more	10.0	1.0
<i>Duration of work in the office</i>		
5 months or less	11.6	14.0
6 - 11 months	12.6	24.0
1 - 2 years	32.2	51.0
3 - 4 years	20.6	10.0
5 - 6 years	11.8	1.0
7 years or more	11.2	-

For maintenance management employees, 51.0 percent of the total respondents have worked in the office for 1 to 2 years, while 24.0 percent have been working for 6 to 11 months and followed by 14.0 percent for 5 months or less. Furthermore, 10.0 percent are maintenance management employees who have worked for a period of 3 to 4 years and only 1.0 percent represent the longest duration of 5 to 6 years. For majority of the respondents, their working duration in the office are based on the duration of service agreement contracts received by the company or organization with the maintenance management of the government offices agencies. Normally, the services duration of a contract agreement is 2 years and would be extended a few year (2 years + 1 year). For maintenance management employees who had served more than three years, this condition occurs based on two factor, firstly, the maintenance management of the company or organization where employees are working successfully secured new contracts with other government offices. The second factor, the employee has received offers to work with a company or organization which recently obtain a new maintenance management contracts.

Before further analysis of a questionnaire survey of Part 2 is done, the Cronbach's Alpha reliability analysis should be done to the data collected. This analysis determines the degree of accuracy and precision of the measurements done by a research instrument. The lower the degree of error for an instrument, the higher reliability of these instruments [15]. In a research, reliability refers to the ability of a study to obtain similar values when the measurement is repeated. In this study, researchers set a measuring instrument that can be considered appropriate and acceptable to be the value of Cronbach's alpha coefficients exceeding 0.70 [16]. Tests results found that the items measuring financial factor have a good reliability value with Cronbach's Alpha value of 0.78 on 11 items measured. Table 6 shows the evaluation of financial factor on the maintenance management for both respondents.

3.1 EVALUATION OF GOVERNMENT OFFICE EMPLOYEES

Table 6 shows, a total of 62.8 percent of respondents agreed and 6.9 percent of respondents strongly agreed that the maintenance management provided financial allocation based on periodic schedule, followed by respondents agreed or strongly agreed by 63.7 percent (58.5 percent agreed, 5.2 percent strongly agreed) that the maintenance management provided financial allocation based on requirement of actual safety and health, while a total of 63.1 percent of respondents agreed and 5.2 percent of respondents strongly agreed that the maintenance management provided financial allocation for preventive maintenance related to safety and health. According to [17], the safety of equipment and employee is improved by preventive maintenance services.

Also, most respondents (71.6 percent) agreed or strongly agreed (66.6 percent agreed, 5.0 percent strongly agreed) that the maintenance management provided financial allocation for corrective maintenance related to safety and health, followed by 65.7 percent of respondents agreed and 4.8 percent of respondents strongly agreed that the maintenance management provided financial allocation for emergency maintenance related to safety and health, while respondents agreed and strongly agreed by 66.5 percent (59.4 percent agreed, 7.1 percent strongly agreed) that the maintenance management provide additional financial allocation if necessary for maintenance related to safety and health.

A total of 80.8 percent of respondents either were neutral, disagreed or strongly disagreed (42.9 percent neutral, 34.0 percent disagreed, 3.9 percent strongly disagreed) that the maintenance management provide yearly financial allocation which often prioritize cosmetic maintenance. Further more, 45.4 percent (41.8 percent agreed, 3.6 percent strongly agreed) of respondents agreed or strongly agreed that the maintenance cost increased especially related to spare part cost. However, there are 64.1 percent (55.9 percent agreed, 8.2 percent strongly agreed) of respondents agreed or strongly agreed that the maintenance management gave more priority to the original spare part although higher costs related to the safety and health.

Table 6. Evaluation of financial factor

No.	Item	Government office employees					Maintenance management employees				
		SA	A	N	D	SD	SA	A	N	D	SD
1.	Maintenance management provide financial allocation based on periodic schedule	6.9	62.8	26.4	3.2	0.7	20.0	62.0	15.0	2.0	1.0
2.	Maintenance management provide financial allocation based on requirement of actual safety and health	5.2	58.5	30.2	5.7	0.4	14.0	64.0	19.0	2.0	1.0
3.	Maintenance management provide financial allocation for preventive maintenance related to safety and health	5.2	63.1	28.3	3.0	0.4	24.0	56.0	17.0	1.0	2.0
4.	Maintenance management provide financial allocation for corrective maintenance related to safety and health	5.0	66.6	24.7	3.2	0.5	24.0	56.0	17.0	2.0	1.0
5.	Maintenance management provide financial allocation for emergency maintenance related to safety and health	4.8	65.7	25.8	3.0	0.7	23.0	59.0	16.0	1.0	1.0
6.	Maintenance management provide additional financial allocation if necessary for maintenance related to safety and health	7.1	59.4	29.9	2.7	0.9	28.0	53.0	15.0	3.0	1.0
7.	Maintenance management provide yearly financial allocation which often prioritize cosmetic maintenance	1.1	18.1	42.9	34.0	3.9	6.0	17.0	35.0	37.0	5.0
8.	Maintenance cost increased especially related to spare part cost	3.6	41.8	33.5	18.8	2.3	24.0	42.0	19.0	12.0	3.0
9.	Maintenance management gave more priority to the original spare part although higher costs related to the safety and health	8.2	55.9	30.9	4.3	0.7	25.0	48.0	21.0	5.0	1.0
10.	Maintenance management often feel that the maintenance is a wasted	1.2	13.4	29.4	47.5	8.5	11.0	16.0	15.0	46.0	12.0
11.	Maintenance management want to earn greater gain and refuse to spend more for maintenance	1.6	13.2	28.3	47.5	9.4	13.0	18.0	22.0	36.0	11.0

Notes: Five-point Likert scale (SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree)

Lastly, a total of 56.0 percent (47.5 percent disagreed, 8.5 percent strongly disagreed) of respondents disagreed or strongly disagreed that the maintenance management often feel that the maintenance is a wasted. Hence, majority of the respondents or 56.9 percent disagreed or strongly disagreed that the maintenance management want to earn greater gain and refuse to spend more for maintenance.

3.2 EVALUATION OF MAINTENANCE MANAGEMENT EMPLOYEES

Based on Table 6, the research found that 82.0 percent of respondent agreed or strongly agreed (62.0

percent agreed, 20 percent strongly agreed) that the maintenance management provided financial allocation based on periodic schedule, followed by 64.0 percent of respondents agreed and 14.0 percent strongly agreed that the maintenance management provided financial allocation based on requirement of actual safety and health, while a total of 56.0 percent of respondents agreed and 24.0 percent of respondents strongly agreed that the maintenance management provided financial allocation for preventive maintenance related to safety and health.

Many respondents (80.0 percent) agreed or strongly agreed (56.0 percent agreed, 24.0 percent strongly agreed) that the maintenance management provided

financial allocation for corrective maintenance related to safety and health (80.0 percent), followed by 59.0 percent agreed and 23.0 percent strongly agreed that the maintenance management provided financial allocation for emergency maintenance related to safety and health, while respondents agreed or strongly agreed by 81.0 percent (53.0 percent agreed, 28.0 percent strongly agreed) that the maintenance management provide additional financial allocation if necessary for maintenance related to safety and health.

For item the maintenance management provide yearly financial allocation which often prioritize cosmetic maintenance, 77.0 percent of the respondents whether neutral, disagreed or strongly disagreed (35.0 percent neutral, 37.0 percent disagreed, 5.0 percent strongly disagreed), followed by 66.0 percent of respondents agreed or strongly agreed (42.0 percent agreed, 24.0 percent strongly agreed) that the maintenance cost increased especially related to spare part cost. This finding supported the research by [18] that the current rate of building repair and maintenance cost in the UK is likely to grow supported by the following facts:-

- total spending on building maintenance in the UK has increased by 66 percent in the last ten years;
- repair and maintenance of building stock represented over 5 percent of Gross Domestic Product, or £36 billion in 1996; and
- repair and maintenance output is expected to increase by 43.6 percent between 1989 and 2001.

This condition does not affect maintenance management, where a total of 73.0 percent of respondents agreed or strongly agreed (48.0 percent agreed, 25.0 percent strongly agreed) that the maintenance management gave more priority to the original spare part although higher costs related to the safety and health. Furthermore, 27.0 percent of respondents agreed or

strongly agreed (16.0 percent agreed, 11.0 percent strongly agreed) that the maintenance management often feel that the maintenance is a wasted. However, according to [19], the primary objective of planned maintenance is the minimization of total cost of inspection and repair, and equipment downtime. Finally, 31.0 percent of respondents agreed or strongly agreed (18.0 percent agreed, 13.0 percent strongly agreed) that the maintenance management want to earn greater gain and refuse to spend more for maintenance.

3.3 DIFFERENCES EVALUATION BETWEEN RESPONDENTS

Table 7 shows evaluation difference between both groups of respondents for items 1 to 6 and item 8 to 9 which is significant at level $p < 0.01$; and item 11 which is significant at level $p < 0.05$.

Maintenance management employees have of mean rank value much higher than the government office workers for all items. This reflects that the maintenance management employees have a more positive evaluation for financial provision made than the government office workers who are affected by the financial provisions. However, the analysis found no significant differences for item 7 and item 10.

IV. CONCLUSION

This paper concluded that maintenance management was aware of safety and health in the maintenance of offices by providing adequate financial provision for these purposes. This research also found that there are significant differences in evaluation between the two respondents against nine financial factors tested, except for items "maintenance management provide yearly financial allocation which often prioritize cosmetic maintenance" and "maintenance management often feel that the maintenance is a wasted"

Table 7. Evaluation between both groups of respondents

Items of financial factor	Mean Rank		Mann-Whitney U	Z
	Government office employees	Maintenance management employees		
1.	321.36	388.51	22399.00**	- 3.77
2.	321.81	385.97	22653.00**	- 3.53
3.	319.69	397.85	21465.00**	- 4.37
4.	320.74	391.98	22052.00**	- 4.07
5.	319.57	398.56	21394.50**	- 4.50
6.	318.17	406.41	20609.00**	- 4.82
7.	331.65	330.64	28014.00	- 0.05
8.	317.56	409.86	20264.50**	- 4.71
9.	322.42	382.54	22996.00**	- 3.22
10.	329.86	340.73	27177.00	- 0.56
11.	323.97	373.80	23870.00*	- 2.56

* p -value < 0.05 , ** p -value < 0.01

REFERENCES

- [1] C. Bahr and K. Lennerts, Quantitative validation of budgeting methods and suggestion of a new calculation method for the determination of maintenance costs, *Journal of Facilities Management*, 8(1), 2010, 47-63.
- [2] A.H.C. Tsang, Strategic dimensions of maintenance management, *Journal of Quality in Maintenance*, 8(1), 2002, 7-39.
- [3] B. Williams, Cost-effective facilities management: a practical approach, *Facilities*, 14(5/6), 1996, 26-38.
- [4] R.M.W. Horner, M.A. El-Haram and A.K. Munns, Building maintenance strategy: a new management approach, *Journal of Quality in Maintenance*, 3(4), 1997, 273-280.
- [5] F. Booty, *Facilities management handbook* (London: LexisNexis, 2003).
- [6] J. Lai J and F. Yik, Expenditure on operation and maintenance service and rental income of commercial buildings, *Facilities* 26(5), 2008, 242-265.
- [7] D. Sherwin, A review of overall models for maintenance management, *Journal of Quality in Maintenance*, 6(3), 2000, 138-164.
- [8] A.S. Ali, Cost decision making in building maintenance practice in Malaysia, *Journal of Facilities Management*, 7(4), 2009, 298-306.
- [9] C.N. Madu, Competing Through Maintenance Strategies, *International Journal of Quality & Reliability Management*, 17, 2000, 937-49.
- [10] A.D. Marco, S. Ruffa and G. Mangano, Strategic factors affecting warehouse maintenance costs, *Journal of Facilities Management*, 8(2), 2010, 104-113.
- [11] A.S. Ali, S.N. Kamaruzzaman, R. Sulaiman and C.P. Yong, Factors affecting housing maintenance cost in Malaysia, *Journal of Facilities Management*, 8(4), 2010, 285-298.
- [12] S. Sliteen, H. Boussabaine and O. Catarina, Benchmarking operation and maintenance costs of French healthcare facilities, *Journal of Facilities Management*, 8(2), 2011, 104-113.
- [13] L.R. Gay, *Educational research: competencies for analysis and application* (Columbus, OH: Charles E. Merrill, 1981).
- [14] M.J. Norusis, *SPSS – statistical data analysis* (Chicago: SPSS Inc., 1990).
- [15] R. Kumar, *Research methodology: a step-by-step guide to beginners* (Sage Publications, 1999).
- [16] J. Nunnally, *Psychometric theory* (New York: McGraw-Hill, 1978).
- [17] S.A. Oke, and O.E. Charles-Owaba, An approach for evaluating preventive maintenance scheduling cost, *International Journal of Quality & Reliability Management*, 23(7), 2006, 847-879.
- [18] M.A. El-Haram and M.W. Horner, Factors affecting housing maintenance cost, *Journal of Quality in Maintenance Engineering*, 8(2), 2002, 115-123.
- [19] M.A. Mirghani, A framework for costing planned maintenance, *Journal of Quality in Maintenance Engineering* 7(3), 2001, 170-182

An overview of e-commerce implementation in developed and developing country; A case study of United State and Nigeria

Lawal Mohammed Ma'aruf & Khadija Abdulkadir

Department Of Mathematicsfaculty Of Sciencesahmadu Bello University. Zaria, +234, Nigeria

Abstract: *Electronic commerce has the potential of improving the efficiency and productivity of any economy. Thus, increase the GDP (Gross Domestic Product) of many countries. However, there has been some doubt about the relevance of e-commerce for developing countries. Currently, there are still a limited number of studies on e-commerce adoption by developing countries. This project is meant to identify factors that could influence proper implementation and adoption of e-commerce in developing countries, particularly Nigeria. Factors affecting the adoption of e-commerce and the condition of Nigeria in relation to e-commerce adoption were also discussed.*

Keywords: *E-commerce, implementation, developed country, developing country.*

I. Introduction

1.1 HISTORY OF E-COMMERCE

Shopping on the internet is certainly a popular past time, an efficient time-saver, and a great way to comparison shop on virtually any kind of item you're interested in. The history of e-commerce as most people think of it has a short but interesting time line. Most people don't realize that e-commerce and its underlying technology have been around for about forty years

The term e-commerce was originally conceived to describe the process of conducting business transactions electronically using technology from the *Electronic Data Interchange* (EDI) and *Electronic Funds Transfer* (EFT). These technologies, which first appeared in the late 1970's, allowed for the exchange of information and the execution of electronic transactions between businesses, typically in the form of electronic purchase orders and invoices. EDI and EFT were the enabling technologies that laid the groundwork for what we now know as e-commerce. The Boston Computer Exchange, a marketplace for used computer equipment started in 1982, was one of the first known examples of e-commerce. Throughout the 1980's, the proliferation of credit cards, ATM machines and telephone banking was the next step in the evolution of electronic commerce. Starting in the early 90's, e-commerce would also include things such as *Enterprise Resource Planning* (ERP), data warehousing and data mining.

It wasn't until 1994 that e-commerce (as we know it today) really began to accelerate with the introduction of security protocols and high speed internet connections such as DSL, allowing for much faster connection speeds and faster online transaction capability. Industry "experts" predicted explosive growth in e-commerce related businesses. In response to these expert opinions, between 1998 and 2000, a substantial number of businesses in Western Europe and the United States built out their first rudimentary e-commerce websites.

The definition of e-commerce began to change in 2000 though, the year of the dot-com collapse when thousands of internet businesses folded. Despite the epic collapse, many of the worlds' most established traditional brick-and-mortar businesses were emboldened with the promise of e-commerce and the prospect of serving a global customer base electronically. The very next year, business to business transactions online became one of the largest forms of e-commerce with over \$700 billion dollars in sales.

1.2 BACKGROUND HISTORY OF THE STUDY

Consumers in e-commerce are faced with a number of risks arising from the general lack of understanding of the operations of the internet. This has been compounded by a number of legal issues which have been largely taken care of in more advanced countries, but which issues are still being grappled with in developing countries such as Nigeria, where internet trading is something fairly new. These issues include the extent to which the communication between the parties is protected (data protection), the formation of a contract on the internet, the legal means of effecting payment in e-commerce, which court will assume jurisdiction in the event of a dispute between parties to an internet contract, and what law or laws will govern the transactions. Is it the law of the seller or that of the buyer or consumer? Other issues relate to cyber crimes that are threatening e-commerce, and also the mode of proving internet-related transactions. An examination of the foregoing issues vis-à-vis the position of the law in developing countries presently is the focus of this project.

1.3 OBJECTIVE OF THE STUDY

The main objective of this study is to enhance the implementation of e-commerce in developing countries, particularly Nigeria. Thus, facilitating international co-operation through trade, creating awareness of e-commerce and its associated benefits, making goods and services available to consumers all over the world irrespective of distance, expansion of the consumer base for manufacturers or producers of goods and services, and a reduction in the costs of service delivery by delivering these electronically.

1.4 JUSTIFICATION OF THE STUDY

Obviously, e-commerce is a vital part of global facilitation policy. Following major initiatives in liberalization in 1991. The need to facilitate international trade through policy and procedural reforms has become the cornerstone of trade and fiscal policies.

Thus, there is need to make a comparative study between these countries in order to provide government and related bodies with adequate information on how to further enhance the implementation of e-commerce in developing countries and also see how far they have been able to implement this fiscal policy with respect to developed countries. Hence the recommendations needed in order to meet up with international standards.

1.5 SIGNIFICANCE OF THE STUDY

The significance of this research project is to facilitate international co-operation through trade, create awareness of e-commerce and its associated benefits, make goods and services available to consumers all over the world irrespective of distance, expands consumer base for manufacturers or producers of goods and services, and also reduce the costs of service delivery by delivering these services electronically.

1.6 SCOPE AND LIMITATION

This study is only interested in the overview of e-commerce implementation between developed and developing Countries. Furthermore, the study is limited to the United States and Nigeria because there is need to compare how electronic commerce is implemented in these two countries.

II. Review of Electronic commerce System

2.1 Introduction

Electronic commerce is a process of buying and selling products, services and information using computer networks and the internet. E-Commerce accelerates global commercial trend by removing the boundaries facing international trade. It benefits from economic advantages such as market expansion, reduction of product source prices, promotion of productivity, reduction of transaction costs and inflation, lowering uncertainty, sharing market information, and aiding in distribution channel efficiency and plays a vital role in an endogenous economic growth. E-Commerce can be a source that improves domestic economic and rapid globalization of production, and development of available technology. Africa and the Middle East suffer from very specific issues that need to be integrated into world agenda and agreements taking place where, the barriers are very well understood and have been researched by many scholars. In the literatures published on e-commerce in developing countries the assumption is that Internet access will enhance e-commerce in these countries, however in most developing countries there are several barriers in expanding e-commerce including basic resources, lack of infrastructure, environmental factors, education and cultural problems. E-commerce can occur within and between three basic participant groups – business, government, and individuals.

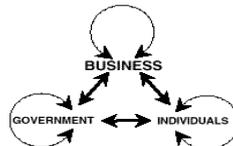


Figure 1: E-commerce participant group

2.2 COMPONENTS OF AN EFFECTIVE E-COMMERCE: Some of the components that you should be looking for (and asking questions about) as you seek a suitable shopping cart to serve your online store include:

SEARCH-ENGINE FRIENDLINESS: You will want your shopping cart solution to provide you with control over your own title tags, meta descriptions and keywords so that you can easily optimize at-will for the search engines. You will also want your solution to automatically generate keyword-rich URLS based on your product and category names and to generate page code that is XHTML 1.0 strict compliant.

PCI COMPLIANT HOSTING AND SECURITY: Because your website will need to accept online payments, you must find a provider who can host your website in a *Payment Card Industry (PCI)* compliant environment. If you don't know anything about PCI Compliance, you will need to understand your role in complying with the DSS (*Data Security Standards*). Currently, the card industry mandates that any merchant, organization or software that processes, stores or disseminates credit card data must be PCI DSS compliant or they risk hefty fines and/or losing the ability to process credit cards altogether thus putting your entire business in jeopardy. Although there is a lot more involved with PCI compliance than just hosting your site in a compliant environment, this is your first step.

UP-SELLING AND CROSS-SELLING FEATURES: Up-selling is offering customers a product that is of greater value than the product they are currently considering or adding options/features to the product they are buying in an attempt to increase their spend. Cross-selling refers to selling items that are complementary to the item they are purchasing. Any suitable shopping cart technology will offer automated features that can accomplish these marketing techniques for you with some simple set-up. Make sure that your shopping cart solution offers these valuable recommendation tools.

PRODUCT REVIEW FEATURE: Product reviews are a very powerful product marketing tool. In fact, research shows that even bad reviews help to sell more product than not showing reviews for a product at all. You will want an e-commerce solution that lets you enable customers to post reviews of the products that you sell. You can always review the posts before you choose to display them or choose to display reviews automatically on the website.

FLEXIBLE PRICING MANAGEMENT: You will want to find a shopping cart that offers flexible pricing features and allows you to set-up multiple pricing options to accommodate things like promotional pricing for any product, customer, order, or time frame, wholesale or retail pricing (or both) displayed to your customers via the same catalog interface and which allows you to set-up affiliate pricing, accepts coupon codes, allows for quantity-based pricing and has tax-exempt support.

ROBUST CATALOG MANAGEMENT: Identify a solution that allows you to get your store started quickly by offering a bulk import/export of your catalog. You will find that this feature will support you throughout the life-cycle of your business and not just at start-up. Make sure that the solution also offers product cloning to simplify catalog additions, automatic resizing of your product images and has some strong inventory management capabilities such as letting you set minimum quantities and identify backordered items.

SOLID ANALYTICS AND REPORTING: Look for a solution that integrates with Google Analytics to demonstrate real-time site and search ranking data and that offers a variety of other product performance reporting capabilities, which you will need to accurately monitor your business. Be sure that you can export and save these reports.

INTEGRATED SHIPPING: You will want your shopping cart solution to integrate with the major shipping providers such as USPS, UPS and Fedex and you will want the ability to markup shipping rate-quotes by a percentage or by a flat amount so that you have the option to charge transparent handling fees.

PRODUCT FEED SUPPORT: You will want your shopping cart solution to have a product feed feature that will allow you to easily create a feed of your products that you can then upload to shopping comparison sites (a.k.a shopping engine) such as Shopzilla, Shopping.com, Google Product Search, etc. and that will allow you to create multiple product feeds for each comparison engine.

GIFT CERTIFICATE/WISHLIST FEATURES: Gift Certificate and Wishlist are two website features that play a major role in holiday sales conversions (according to a Demandware study). 64% of those surveyed said gift certificates equaled 25% of their online purchases and half of those surveyed said they shopped using someone else's wishlist. So, be sure to utilize these features on your site and do not consider an ecommerce solution that does not boast a robust online gift certificate and wishlist feature-set.

2.3.0FEATURES OF E-COMMERCE

Nowadays E-Commerce is a global word. Although it is omnipresent but we never realize its importance primarily because it is known by different names. People do site promotion, SEO, affiliate marketing, and many other things but the goal is same i.e. to get clients and sell the products or services of the company. 'E' is just a medium to transact online. The following are some peculiar features of e-commerce which makes it considerably appreciable.

UBIQUITY: In traditional commerce, a marketplace is a physical place we visit in order to transact. For example, television and radio are typically directed to motivating the customer to go someplace to make a purchase. E-commerce is ubiquitous, meaning that it is available just about everywhere at all times. It liberates the market from being restricted to a physical space and makes it possible to shop from your desktop. The result is called a market space. From consumer point of view, ubiquity reduces transaction costs - the cost of participating in a market. To transact, it is no longer necessary that you spend time and money traveling to a market. At a broader level, the ubiquity of e-commerce lowers the cognitive energy required to complete a task.

GLOBAL REACH: E-commerce technology permits commercial transactions to cross cultural and national boundaries far more conveniently and effectively as compared to traditional commerce. As a result, the potential market size for e-commerce merchants is roughly equal to the size of world's online population.

UNIVERSAL STANDARDS: One strikingly unusual feature of e-commerce technologies is the technical standards of the Internet and therefore the technical standards for conducting e-commerce are universal standards, that is, they are shared by all the nations around the world.

INTERACTIVITY: Unlike any of the commercial technologies of the twentieth century, with the possible exception of the telephone, e-commerce technologies are interactive, meaning they allow for two-way communication between merchants and consumer.

INFORMATION DENSITY AND RICHNESS: The Internet vastly increases information density. It is the total amount and quality of information available to all market participants, consumers and merchants. E-commerce technologies reduce information collection, storage, communication and processing costs. At the same time, these technologies increase greatly the accuracy and timeliness of information, making information more useful and important than ever. As a result, information becomes plentiful, cheaper and of higher quality. Information richness refers to the complexity and content of a message.

PERSONALIZATION: E-commerce technologies permit personalization. Merchants can target their marketing messages to specific individuals by adjusting the message to a person's name, interests and past purchases. The technology also permits customization. Merchants can change the product or service based on user's preferences or prior behavior.

2.4 IMPORTANCE OF E-COMMERCE

Electronic commerce, or e-commerce, which literally means business trading through the Internet, has been around the globe since mid 90s. However, until the recent few years, e-commerce is getting more and more attention from entrepreneur and consumers, both local and international. One of the main reasons is due to the highly successful operations of some well known names on the Internet, such as eBay, Yahoo and Dell. The sales revenue these companies show in their annual reports are without doubt, one of the biggest factors why e-commerce is important in the commercial market nowadays.

E-commerce proved its importance based on the fact where time is essence. In the commercial markets, time plays an important role to both the business and consumers. From the business perspective, with less time spent during each transaction, more transaction can be achieved on the same day. As for the consumer, they will save up more time during their transaction. Because of this, e-commerce steps in and replaced the traditional commerce method where a single transaction can cost both parties a lot of valuable time. With just a few clicks in minutes, a transaction or an order can be placed and completed via the internet with ease. For instance, a banking transaction can be completed through the Internet within a few minutes compared to the traditional banking method which may take up to hours. This fact obviously proves that e-commerce is beneficial to both business and consumer wise, as payment and documentations can be completed with greater efficiency.

From the business viewpoint, e-commerce is much more cost effective compared to traditional commerce method. This is due to the fact where through e-commerce, the cost for the middleperson to sell their products can be saved and diverted to another aspect of their business. One example is the giant computer enterprise, Dell, which practice such a method by running most of their business through internet without involving any third parties. Aside from that, marketing for e-commerce can achieve a better customer to cost ratio as putting an advertisement on the internet is comparably much cheaper than putting up a roadside banner or filming a television commercial. For e-commerce, the total overheads needed to run the business is significantly much less compared to the traditional commerce method. The reason due to that is where most of the cost can be reduced in e-commerce. For example, in running an e-commerce business, only a head office is needed rather than a head office with a few branches to run the business. In addition to that, most of the cost for staff, maintenance, communications and office rental can be substitute by a single cost, web hosting for the e-commerce business.

To both the consumers and business, connectivity plays an important part as it is the key factor determining the whole business. From the business point of view, e-commerce provides better connectivity for its potential customer as their respective website can be accessed virtually from anywhere through Internet. This way, more potential customers can get in touch with the company's business and thus, eliminating the limits of geographical location. From the customer standpoint, e-commerce is much more convenient as they can browse through a whole directories of catalogues without any hassle, compare prices between products, buying from another country and on top of that, they can do it while at home or at work, without any necessity to move a single inch from their chair. Besides that, for both consumers and business, e-commerce proves to be more convenient as online trading has less red tape compared to traditional commerce method. In global market sense, the appearance of e-commerce as a pioneer has opened up various windows of opportunities for a variety of other companies and investors. For instance, due to the booming of e-commerce, more and more resources are being directed into electronic securities, internet facilities, business plans and new technologies. In result of this phenomenon, a variety of new markets have emerged from e-commerce itself giving a boost to the global market.

2.5 BENEFITS OF E-COMMERCE

The impact of e-commerce to businesses can be huge. E-commerce can transform the way products and services are created, sold and delivered to the customer. It can also change the way in which a company works with its partners. Some benefits of e-commerce are as follows.

IMPROVED PRODUCTIVITY: Using e-commerce, the time required to create, transfer and process a business transaction between trading partners is significantly reduced. Furthermore, human errors and other problems like duplications of records are largely eliminated with the reduction of data-entry and re-entry in the process. This improvement in speed and accuracy, plus the easier access to document and information, will result in increase in productivity.

COST SAVINGS: According to a 1999 report by Giga information Group, conducting business online will save companies around the world an estimated USD1.25 trillion by 2002. This compares to total savings of USD 17.6 billion in 1998-The cost savings stem from efficient communication, quicker turnaround time and closer access to markets.

BETTER CUSTOMER SERVICE: With e-commerce, there is better and more efficient communication with customers. In addition, customers can also enjoy the convenience of shopping at any hour, anywhere in the world.

CONVENIENCE: E-commerce is very convenient to the consumers since the sites operates 24 hours a day 7 days a week, thus, allowing transaction to be done at any time.

REDUCED ERRORS: The automated process tends to produce fewer errors than the traditional process since it is done electronically.

UNLIMITED SHELF SPACE: E-commerce companies can display/shelve an unlimited number of goods from which the consumers can choose whatever they want.

INCREASED GLOBAL PRESENCE: E-commerce companies can be accessed by people from all over the world, that is, one can purchase from wherever he/she is.

2.6 ADOPTION AND IMPLEMENTATION OF ELECTRONIC COMMERCE

It is argued that diffusion of innovation theory is relevant to the study of e-commerce, and that e-commerce has unique features suggesting that e-commerce needs its own specific study. E-commerce has technical component similar to other IT innovations, but e-commerce also has inter organizational elements which distinguish it from other types of innovations.

According to Rogers (1983), diffusion is the process during which an innovation is communicated through certain channels over time among members of a social system. However, prior to that, the decision has to be made on whether or not the organization should adopt a new innovation or practice for the business. Rogers distinguishes diffusion from adoption by stating that adoption is a decision to make full use of an innovation as the best course of action, whereas rejection is a decision not to adopt an available innovation (Rogers, 1983:21). In this study, e-commerce adoption is therefore defined as the decision to make use of e-commerce to conduct business or transaction with its trading partners and e-commerce implementation is taking the necessary actions to facilitate and execute e-commerce into the business practice or process. These definitions make it clear that e-commerce implementation involves the use of computer and telecommunications technologies to improve business processes. In many organizations, however, implementing e-commerce requires management and staff to deal with quite new technologies, applications and processes. A whole range of technological, management and business issues arise during the implementation and need to be addressed. For an effective e-commerce implementation, some steps need to follow.



Figure 2: Steps to effective e-commerce implementation

2.7 FACTORS THAT CONTRIBUTES TO E-COMMERCE IMPLEMENTATION SUCCESS

From the many variables that might be considered as indicators for success of an Internet-based e-commerce, satisfaction has been chosen in this study. In the IS literature, satisfaction is one of the variables most often used as a success indicator (Crum, *et al.*, 1996; Delone and McLean, 1992; Alavi and Joachimsthaler, 1992; Seddon and Kiew, 1994; Thong and Yap, 1996; Wierenga and Ophuis, 1997). Being one of the popular constructs of IS implementation success, satisfaction has been found to correlate with usage (Baroudi, *et al.*, 1996; Gelderman, 1995; Wierenga and Ophuis, 1997; Zinkhan *et al.*, 1987). Other scholars have even suggested that low satisfaction causes discontinuity in the usage of information technology (Evans, 1976; Thong and Yap, 1996). Hence, satisfaction is employed as a construct of e-commerce implementation success in this study.

E-commerce offers huge potential to SMEs, including potential strategic benefits such as possibilities of creating new industries, developing new content and chances to find or create employment.

commerce. In addition, developing countries often have different cultures and business philosophies, which limit the applicability and transferability of the e-commerce models designed by Western countries.

Internet usage in Nigeria is still relatively low compared to developed countries and e-commerce is still in an elementary stage, notwithstanding there is growing awareness in of the benefits and opportunities offered by e-commerce amongst Nigerians (Bamodu, 2005).

E-commerce activity in Nigeria is steadily growing as a result of vast improvements in telecommunication services. The Telecommunications Industry has experienced exponential growth in the last four years with about 20 million telephone lines connected to date (Ndukwe, 2006). Alongside this growth, there has been an increase in the number of private telephone operators offering fixed wireless service with data transfer capabilities leading to an increasing number of people with internet access at home in the major cities and in some rural areas. Moreover, with the reduction in tariffs and further cuts expected, telecommunications service will become more affordable and essential to many Nigerians.

The recent advances in the telecoms market, and the explosion in the number of subscribers, demonstrate the potential market for information communication technology services generally in Nigeria (Ndukwe, 2006). Given Nigeria's sizable population it is a potentially lucrative market for electronic commerce services.

According to the Economist Intelligence Unit (EIU, 2006), the stock of personal computers (PCs) per 1,000 persons grew from 10.66 in 2004 to 11.09 in 2005. The greatest obstacle to the growth of e-commerce is low PC penetration. However affordable Asian technologies and falling microchip prices have fuelled a market in cloned as well as branded PCs. Electronic banking is one area of e-commerce that has proven successful in Nigeria (EIU, 2006). Nigerian banks are increasingly seeking to provide general banking facilities online. Internet banking is slowly and steadily gaining ground, banks have set up websites which publish corporate information and allow customers to carry out some form of transaction – limited in most cases. Despite the growing focus on internet banking not all banks are moving at the same pace, some still have only informational websites. Given that the banks exist in the same operational environments, some other factors other than the often cited country context must be responsible for the difference in attitude to electronic commerce among banks.

The initial advances that have been made in electronic banking is a step in the right direction and could be a motivator in the adoption of e-commerce services amongst Nigerians (EIU, 2006).

REASONS FOR SLOW DEVELOPMENT OF E-COMMERCE IN DEVELOPING COUNTRIES: One key reason for the slow development of e-commerce in developing countries is that there is not an overall policy framework covering aspects such as technical, economic and political. Policy considerations when creating an enabling environment include:

Encryption and Decryption techniques provide authentication, authorization, confidentiality and integrity to services, increasing the security of e-commerce transactions. They are necessary, for instance, for processing credit card information.

Digital signatures and electronic contracts are relevant, for instance in cases of dispute between trading partners in an e-commerce transaction.

Certification authorities secure electronic transactions and act as trusted third parties to verify information about parties. African certification authorities must take part in the international framework for supporting ways to link certification mechanisms and the mutual recognition of different certification authorities.

Consumer protection: In an electronic market place it is not easy for consumers to identify and localize suppliers so it is necessary to promote protection mechanisms.

Electronic payments: Online payment using credit cards is a missing component of the African business environment, which is often cash-based. Electronic payments will involve central banks and other trade and financial institutions.

Copyright and intellectual property rights: Legislation on copyright and intellectual property rights on the Internet is still in its infancy, and uncertainty about such legislation contributes to inhibiting business investment.

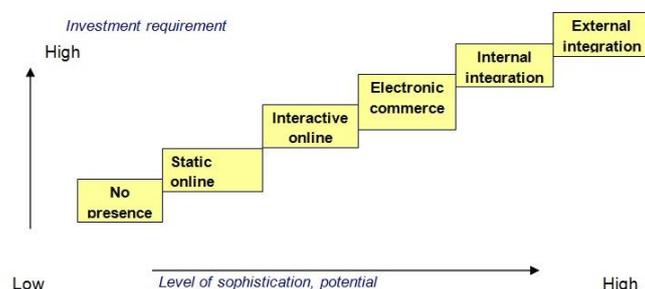


Figure 4: E-commerce stages of growth in

2.10 FACTORS AFFECTING E-COMMERCE IMPLEMENTATION

Some of the factors affecting the adoption of e-commerce in developing countries are:

- Perceived complexity
- Perceived benefits
- Organizational competence
- Perceived compatibility
- Supporting industries e-readiness
- Management support
- Market e-readiness
- IT capability
- Government e-readiness

PERCEIVED COMPLEXITY

Perceived complexity refers to the degree to which an innovation is perceived as difficult to understand and use. New ideas that are simpler to understand are adopted faster than those requiring the adopter to develop new skills and understanding (Rogers, 1995). Akbulut (2002) state that the complexity of a technology has a major effect on the adoption decision, while Chwelos et al. (2002) state that complexity is a strong inhibitor of intent to adopt innovation.

PERCEIVED BENEFITS

Perceived benefits refer to the extent of managements recognition of the relative advantage of adopting e-commerce to the organization. A perceived benefit is an important factor in adoption of new innovations (Iacovou et al., 1995; Rogers, 1995). Rogers (1995) defined Relative advantage as the extent to which an innovation is perceived as better than the idea it supersedes or its nearest alternative. Relative advantage can be measured in financial terms; however, social status, comfort, and satisfaction are important factors as well. The amount of objective advantage of an innovation has a great effect, what affects adoption of an innovation is whether the innovation is viewed as advantageous. The greater the perceived relative advantage of an innovation, the more rapid its rate of adoption will be (Rogers, 1995).

ORGANIZATIONAL COMPETENCY

The availability of employees with competency for producing new ideas is important for e-commerce adoption (Mohr, 1969). Organizational competency refers to the availability of employees with adequate experience and exposure to information and communication technology and other skills (such as business strategy) that are needed to adequately staff e-commerce projects (Molla & Licker, 2005).

PERCEIVED COMPATIBILITY

Perceived compatibility refers to the degree to which an innovation is perceived as being consistent with existing needs, values, past experiences, and technological infrastructure of potential adopters (Rogers, 1995 & Rogers 1983). An innovation might be perceived as technically or financially superior in accomplishing a given task, but it may not be adopted, if a potential adopter views it as irrelevant to its needs (Rogers, 1995).

SUPPORTING INDUSTRIES E-READINESS

Supporting Industries e-readiness refers to “the assessment of presence, development, service level and cost structure of support-giving institutions such as telecommunications, financial, trust enablers and the IT industry, whose activities might affect the e-commerce initiative of businesses in developing countries” (Molla & Licker, 2005). Existence of adequate IT infrastructure is a necessary condition for the take-off of and development of e-commerce (Palacios, 2003); since organizations would rather concentrate on their core competencies, it is vital that there are other organizations whose main activity is provision of IT infrastructure and services.

MANAGEMENT SUPPORT

Top management support has been identified as crucial in the acquisition and diffusion of innovation (Orlikowski, 1993). Top management consists of individuals with power and authority to make strategic decisions; thus they can develop a clear-cut e-commerce vision and strategy while at the same time sending signals to different parts of the organization about the importance of e-commerce. Given the limited nature of organizational resources and the many competing projects, top management support ensures that an e-commerce innovation project will get the required resources and capabilities.

MARKET E-READINESS

Market e-readiness refers to “the assessment that an organization’s business partners such as customers and suppliers allow an electronic conduct of business” (Molla & Licker, 2005). for ecommerce to thrive sellers and buyers have to be willing to exchange goods and services for money online (Turban, 2004). Thus, an organization considering adoption may first examine the willingness of its existing customers and suppliers to do business online or the likelihood of generating new business online.

I.T CAPABILITY

IT capability refers to the level of IT resources and personnel IT knowledge of an organization (Akbulut, 2002). Access to adequate equipment in the organization is a major determinant of the adoption of new technologies (Newcomer and Caudle, 1991). Cohen & Levinthal (1990) state that an organization’s ability to appreciate an innovation, to assimilate it, and apply it to new ways is largely a result of the firms preexisting knowledge in areas relating to the intended innovation.

GOVERNMENT E-READINESS

Government e-readiness refers to “the organizations’ assessment of the preparation of the nation state and its contributions to promote, support, facilitate and regulate ecommerce and its various requirements” (Molla & Licker, 2005). The government has a strong role in promoting and spreading the benefits of electronic commerce (Bandyo-padhay, 2002). Governments can provide an enabling environment in which ecommerce can realize its full potential. They can help address the problems and challenges of awareness, infrastructure develops, local content creation depending on languages used and cultures prevailing in the local environment (Kamel, 2006).

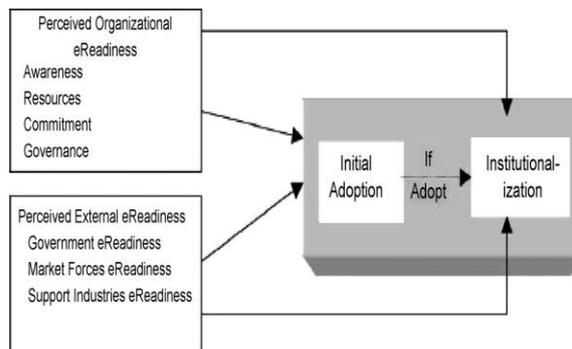


Figure 5: Factors affecting e-commerce implementation

III. E-Commerce Implementation In United States

When the United States sneezes, an economists' proverb says, the rest of the world catches a cold. The first ingredient of a nation's economic system is its natural resources. The United States is rich in mineral resources and fertile farm soil, and it is blessed with a moderate climate. The second ingredient is labor, which converts natural resources into goods. The number of available workers and, more importantly, their productivity help determine the health of an economy. Labor-force quality continues to be an important issue. Today, Americans consider "human capital" a key to success in numerous modern, high-technology industries. As a result, government leaders and business officials increasingly stress the importance of education and training to develop workers with the kind of nimble minds and adaptable skills needed in new industries such as computers and telecommunications.

The United States is said to have a mixed economy because privately owned businesses and government both play important roles. The American free enterprise system emphasizes private ownership. Private businesses produce most goods and services, and almost two-thirds of the nation's total economic output goes to individuals for personal use (the remaining one-third is bought by government and business). The consumer role is so great, in fact, that the nation is sometimes characterized as having a "consumer economy."

This emphasis on private ownership arises, in part, from American beliefs about personal freedom. From the time the nation was created, Americans have feared excessive government power, and they have sought to limit government's authority over individuals -- including its role in the economic realm. In addition, Americans generally believe that an economy characterized by private ownership is likely to operate more efficiently than one with substantial government ownership. The United States accounts for about 80 per cent of the global total of electronic commerce. While its share will probably decrease, it does not face some of the constraints that confront Europe and Asia, such as high cost and the lack of sufficient bandwidth and the slow pace of planned liberalization of the telecommunications sector.



Figure 6: E-commerce growth rate in United State between 1999 & - commerce has emerged as the fastest growing sector of the U.S. marketplace. Despite the contraction in the high-tech industry during the

recent recession, firms have continued to enter and expand their presence in e-commerce, and consumers have increased the number of purchases made online. E-commerce currently represents a very small share of overall commerce, but it is expected to continue to expand rapidly in coming years. As e-commerce grows, so will its impact on the overall economy. The primary route by which e-commerce will affect the economy at large is through its impact on productivity and inflation. Business and consumers that use e-commerce benefit from a reduction in costs in terms of the time and effort required to reach for goods and services and to complete transactions. This reduction in costs results in higher productivity. An even larger increase in economy – wide productivity levels may result from productivity gains by firms not engaged in e-commerce as they respond to this new source of competition.

Continued expansion of e-commerce may also lead to downward pressure on inflation through greater competition, cost savings, and changes in price-setting behavior of sellers. It was examined that economic factors that have contributed to the rapid growth of e-commerce and assesses how the future growth of e-commerce may affect the overall economy. That concluded e-commerce continues to grow rapidly, it could lead to an increase in productivity growth and downward inflationary pressures that persist for several years.

WHAT CONTRIBUTED TO EFFECTIVE E-COMMERCE IN UNITED STATES: In the United States, some electronic commerce activities are regulated by the *Federal Trade Commission* (FTC). These activities include the use of commercial e-mails, online advertising and consumer privacy. The CAN-SPAM Act of 2003 establishes national standards for direct marketing over e-mail. The Federal Trade Commission Act regulates all forms of advertising, including online advertising, and states that advertising must be truthful and non-deceptive. Using its authority under Section 5 of the FTC Act, which prohibits unfair or deceptive practices, the FTC has brought a number of cases to enforce the promises in corporate privacy statements, including promises about the security of consumers' personal information. As result, any corporate privacy policy related to e-commerce activity may be subject to enforcement by the FTC.

E-COMMERCE IMPLEMENTATION IN NIGERIA: The economy of Nigeria is a middle income, mixed economy emerging market with well-developed financial, legal, communications, transport, and entertainment sectors. It is ranked 31st in the world in terms of GDP (PPP) as of 2009, and its emergent, though currently underperforming manufacturing sector is the second-largest on the continent, producing a large proportion of goods and services for the West African region.

Previously hindered by years of mismanagement, economic reforms of the past decade have put Nigeria back on track towards achieving its full economic potential. Nigerian GDP at purchasing power parity more than doubled from \$170.7 billion in 2005 to \$374.3 billion in 2010, although estimates of the size of the informal sector (which is not included in official figures) put the actual numbers closer to \$520 billion. Correspondingly, the GDP per capita doubled from \$1200 per person in 2005 to an estimated \$2,500 per person in 2009 (again, with the inclusion of the informal sector, it is estimated that GDP per capita hovers around \$3,500 per person). It is the largest economy in the West Africa Region, 3rd largest economy in Africa (behind South Africa and Egypt), and on track to becoming one of the top 30 economies in the world in the early part of 2011.

In the past few years, Nigerian banks and the financial services industry in general, have embraced e-banking/e-money made possible by advancements in information-based technology. I do not think they have a choice, anyway. In the competitive environment of this twenty-first century, technology is "how to beat the other guys." Indeed, technological innovation presents banks with an opportunity to gain a competitive advantage through cost-effective delivery system and to use these systems to generate fee income.

NIGERIAN EXPERIENCE IN E-COMMERCE: Currently, the concept of e-payment/e-money in Nigeria is card-based and involves in the main, only prepaid cards. The CBN welcomes, and has indeed been very supportive of the introduction and usage of such e-money products in the country, as this could, in principle, improve efficiency in payment operations. The introduction of such e-payment products in Nigeria commenced in 1996 when the CBN granted Allstates Trust Bank approval to introduce a closed system electronic purse called ESCA. This was followed in February 1997, with the introduction of a similar product called "Paycard", by Diamond Bank. The card based e-money products assumed an open platform with the authorisation in February 1998, of Smartcard Nigeria Pie, a company floated by a consortium of 19 banks to produce and manage cards issued by the member banks. Another consortium of more than 20 banks under the auspices of Gemcard Nigeria Limited obtained CBN approval in November 1999 to introduce the "Smartpay" scheme. The number of participating banks in each of the two schemes had since risen to over 35 as at July 2002.

The CBN has additionally granted approval to a number of banks to introduce international money transfer products, telephone banking and on-line banking via the Internet, though on a limited scale. Mention must also be made of the deployment of automatic teller machines (ATM) by a few banks to facilitate cards usage and further enhance their service delivery. We are also quite aware that virtually all Banks in Nigeria now has a website. Indeed, a number of these sites has capability that supports and actually pen-nits the conduct of e-banking, a subset of e-commerce. We anticipate that very soon, many more banks will enhance their website capability to support online transaction processing and electronic bill paying services.

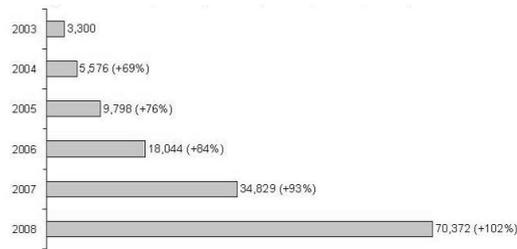


Figure 7: E-commerce growth rate in Nigeria between 2003-2008

OBSTACLES FACING E-COMMERCE IMPLEMENTATION IN NIGERIA: Some of the obstacles facing Nigerian e-commerce and which has contributed to the growth of e-commerce in the United States are:

- the high cost of equipment, integration and connectivity;
- the lack of awareness of what electronic commerce is;
- the lack of legal framework conducive to e-commerce;
- the lack of skilled personnel;
- the uncertainty of the bank regulations

HIGH COST OF EQUIPMENT

Nigerians lag far behind developed country markets in the availability of the technical pre-requisites for conducting electronic commerce. The gaps in the two main requirements for Internet that are telephone and computer availability highlight the difference. For instance, 65 per cent of households in the world have no telephone, whereas 90 per cent of households in high-income countries have a telephone (ILO, 2000).

The personal computer ratio per 100 inhabitants is 18% for high-income countries, 2.3% for medium-income nations and just 0.1% for low-income countries to which ECOWAS member states belong (ILO, 2000). In the United States, at least one out of three persons uses the Internet, compared to only one in every 10,000 in developing countries. Teledensity (main lines per 100 inhabitants) is 48% for developed countries, 10% for middle income and 1.5% for the least developed countries (LDCs) among which the vast majority of ECOWAS nations.

Furthermore, in West African countries telecommunications services are often unreliable, high cost or both. There are also enormous differences in access to telecommunications both between and within countries. For instance, while in West Africa a considerable proportion and sometimes the majority of the population lives in rural areas, over 80 per cent of the main telephone lines are located in urban areas.

AWARENESS ABOUT E-COMMERCE

As far as user awareness is concerned, the successful experience in other parts of the world may help in formulating policies to foster electronic commerce. In this regard, trade associations and chambers of commerce, together with governments, can be good vehicles to promote the use of electronic commerce through awareness campaigns including the organization of “information days”, multimedia demonstrations and the production of information brochure aimed at the general public (Wilson, 1998).

LEGAL AND REGULATORY FRAMEWORK

Another critical issue that West African states need to address is how to make their legal framework conducive to e-commerce transactions. This is important because the creation and adoption of a responsive policy and regulatory framework would stimulate the rapid development of e-commerce in the region, with attendant benefits for the economies and citizens of ECOWAS. It would also have the advantage of clarifying rules and removing some of the barriers set by existing laws that are now rendered anachronistic by modern technological developments. Finally, it would help settle the issues of thrust and abuses in electronic transactions.

The absence of legacy e-commerce legislations at the individual member state level provides an opportunity for ECOWAS and WAEMU to jump-start the process by creating model of e-commerce legislation that can be implemented without the extensive harmonization efforts that would be required if legislation existed in one or more member countries. The sub-regional policy and legislation would also have the advantage of being cost-effective as a single undertaking to be extended and adapted to individual nations. This would also permit efficient capacity building both at the RECs and national levels, to ensure that there is the requisite knowledge among the parties responsible for policy making, legislation and implementation.

HUMAN RESOURCE

Electronic commerce is progressively changing the way businesses are done. The world have moved from an industrial economy where machines dominated productivity, to an information-based economy where intellectual content is the dominant source of value added and which knows no geographic boundaries. In this new environment, education and continuous learning will be essential not only for managers and workers but for all. As industry, commerce, and services are transformed by technology, many skills need to be improved or acquired.

If Nigerians are to benefit from this new technological and economic boom that the growth of e-Commerce represents, they will need to have the most important component in place, the human resource. Today's knowledge revolution largely depends on human capital. Some developing countries and some sections of most developing countries in

Nigeria have this in abundance and opportunities of the new economy could provide them with excellent results (ILO, 2000). On the other hand, there are many countries mainly in Africa that are far behind in this area and therefore need very proactive policies and investment in education (especially technical) to realize the potential. After all to be e-literate, citizens first need to be literate.

As e-Commerce develops and the more advanced stages of commercial exchange (contracting, payment reconciliation and auditing) are carried out electronically, more specific skills are going to be required. Even for surfing the Internet for a product or service, basic familiarity with the computer and knowledge of the Internet is needed. Moreover, extensive language knowledge may be an additional requirement if foreign Internet sites are to be browsed. From website design, to electronic credit management and software and hardware maintenance, all require skills that may not be so easily available in several West African countries. Capacity building in the field of information technology, in the knowledge of the existence of a regional and global market for such skills, is therefore crucial. The development of electronic commerce puts a premium on the development of education and training policies, to ensure that training institutions' curricula meet with the needs of industry.

The training needs should firstly focus on activities in the area of human resource development for electronic commerce, in particular through the training of trainers. Three main targets could be focused upon initially, namely:

- Policy makers and negotiators;
- Small and medium-sized enterprise managers and trade practitioners;
- Technicians and workers, including specialists in electronic data interchange and the Internet. However building the human resource goes far beyond training. It is a continuous learning process. For West Africa, the process should begin with sensitization as well as web design and development of commercial applications. Governments should take the initiative to then pass it on to the private sector. Here the state could pursue its efforts in building partnerships with civil society in order to offer proper training tools in the area of electronic commerce. Similarly international organizations such as UNECA, UNDP, WTO also have a responsibility to assist countries in this crucial area of advocacy and human resource development for this new technology¹. Several of these organizations are in fact working towards this and therefore collaboration and common programmes with countries would be beneficial for the sub-region as a whole.

FINANCIAL FRAMEWORK

In order to encourage e-Commerce development as a support to regional integration in West Africa, banking procedures within the sub-region need also to be aligned and made compatible with digital trade transactions, so that consumers and buyers can easily make purchases locally through the Internet. The higher the development of electronic connections within the region, the higher will be the purchases from within. Along these changes, e-commerce would also require a financial and banking framework that allows for electronic payments and transfers. This would include requirements for certification of documents, electronic signatures, confidentiality and privacy. Therefore West African countries will need to put in place an electronic network (between financial institutions) in addition to the legal framework to allow for such transactions. Banking laws and regulations thus need to be adjusted to the new formats and requirements.

IV. Summary, Conclusion, And Recommendation

3.1 SUMMARY

Electronic commerce has the potential to improve efficiency and productivity in many areas and, therefore, has received significant attention in many countries. However, there has been some doubt about the relevance of e-commerce for developing countries. Currently, there are still a limited number of studies on e-commerce implementation by developing countries. To address the gap in the literature, this study has assessed the implementation of e-commerce in the United States, as an example of a developed country and also Nigeria, as an example of a developing country. Some factors affecting the adoption and implementation of e-commerce and the condition of Nigeria in relation to e-commerce implementation are discussed in this project.

3.2 RECOMMENDATION

The provided recommendations of this research regarding the comparative study of e-commerce implementation in order to ensure e-commerce success in developing countries is to:

- Address high cost of equipment, integration and connectivity;
- create awareness of what electronic commerce is;
- create a legal framework conducive to e-commerce;
- provide skilled personnel;
- address financial and bank regulatory issues.

V. CONCLUSION

Based on the study conducted, its obvious, that e-commerce is indeed relevant to Nigeria, despite the current limitations with the existing infrastructure and other issues related to the economical and socio-cultural conditions. E-commerce will change the mode of doing business, thus offer excellent opportunities for growth in developing nations, particularly in Nigeria.

References

- [1] Aida O.M.; Assefa B. and Sizo M. “ E-commerce Challenges in Africa: Issues, Constraints, Opportunities”.(2003)
- [2] Aitken, I. “Value-driven I.T management: Commercializing the IT function,” (2003)Alev M.E.; Vincent F. Y. and William L. M. “E-commerce in Developing Countries: Issues and Influences.” University of San Francisco. Press (2002)
- [3] Angehm, A. “Designing mature internet business strategies”: the ICDT model. European management Journal, vol. 15(4), pp.361-369. (1997).
- [4] Ann, L.F. “The impact of Electronic Commerce on Business-Level Strategies”, Journal of Electronic Commerce Research, University of Nebraska – Omaha. Vol. 1, No. 1, pp. 13. (2000)
- [5] Buhrmann, A.. “E-bussiness in the year of the gap”. Bussiness Vol.2.0:pp. 48-59, (2002)
- [6] Concord, M.A.” Trends in Electronic Procurement and Electronic Commerce and Their Impact on Small Business”, by Innovation & Information Consultants, Inc. (2004).
- [7] Imala, O.I. “Electronic Commerce and Telecommunications in Nigeria: Bank regulator perspective.”
- [8] Lavin, A. “Factors Affecting Ecommerce Adoption in Nigerian Banks.”
- [9] Udo, G.J. “Factors Affecting E-commerce Website Effectiveness”, The Journal of Computer Information Systems.

Strategies in Energy Conservation and Management of Heating, Ventilating and Air-Conditioning (HVAC) System of a Data Center

Shrikant Belsare

M.E. Mech. – Heat Power, Rajarshi Shahu College of Engineering, Pune, India.

ABSTRACT: This paper addresses the issue of energy performance of data centers (Server Farms). The primary objective of the study is to understand an empirical energy use pattern of data centers under various conditions like tropical climatic conditions, thermal management problems (HVAC) which consumes about 60 – 70 % of the energy, and suggest strategies and give guidance for data centers’ design, operation and maintenance and retrofitting to achieve better energy performance. Identification of different zones in a data center, actual energy use characteristics, the airflow and temperature patterns, cooling of electronic equipments, chilled air distribution, lightening system, design criteria, and energy and cost saving potentials were analyzed and compared. Methodology of energy performance evaluation of data centers was discussed. The study concludes that data centers were high energy consuming areas, substantially the HVAC area. Power demands were often grossly over-provided in these facilities. This led to substantial increase in capital, running cost, harmful environmental effects. Approximately 40% of energy consumption could be conserved through monitoring, evaluation of implemented conservation measures, efficient designs of base infrastructure and energy consuming systems, practical benchmarking, latest technologies and better practices.

I. INTRODUCTION

A **data center** or **datacenter** (or **datacentre**), also called a server farm, is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls, which includes air conditioning, fire suppression and security devices.

II. APPLICATION AND COMPONENTS OF DATA CENTERS

The main purpose of a data center is running the applications that handle the core business and operational data of the organization. Often these applications will be composed of multiple hosts, each running a single component. Common components of such applications are databases, file servers, application servers, middleware, and various others. Data centers are also used for off site backups. In short, data center components include Compute servers, Communication devices, Tape storage, Storage servers, Workstations (standalone) and other rack- and cabinet-mounted equipments.

Five Major Functional Components of a Data Center are: -

- => Data Center Equipments Area => Data Center Cooling system
- => Data Center Electrical System => Data Center Fire Protection System
- => Data Center Physical Security

III. REQUIREMENTS OF DATA CENTER

The Telecommunications Industry Association (TIA) in April 2005 responded with the TIA-942 Telecommunications Infrastructure Standards for Data Centers, the first standard to specifically address data center infrastructure.

- Site space and layout
- Cabling infrastructure
- Tiered reliability
- Environmental considerations

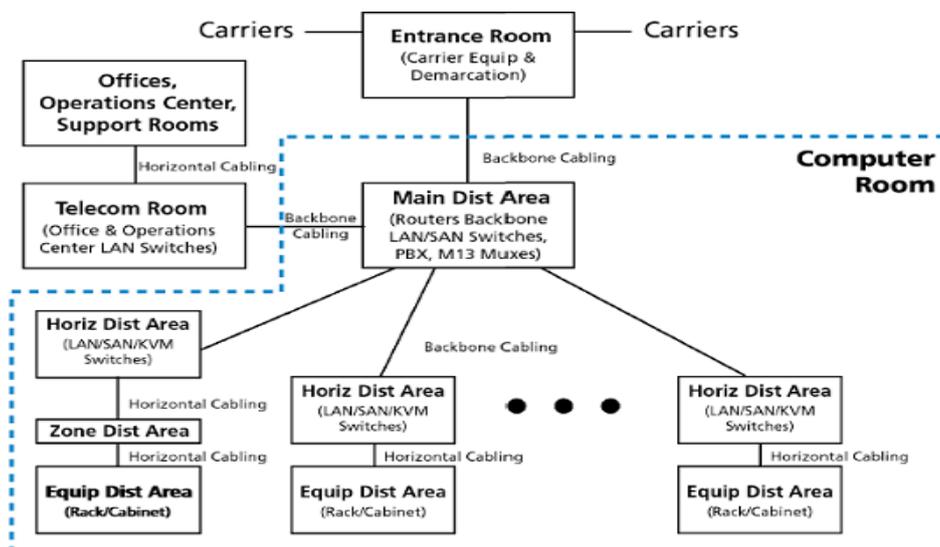


Figure I. Data Center Layout as per TIA –942 standards

According to TIA-942, a data center should include the key functional areas: -

- => Main Distribution Area (MDA)
- => Zone Distribution Area (ZDA)
- => Cabling Infrastructure
- => Equipment Distribution Area (EDA)
- => Backbone and Horizontal Cabling

IV. TOP DATA CENTER CHALLENGES

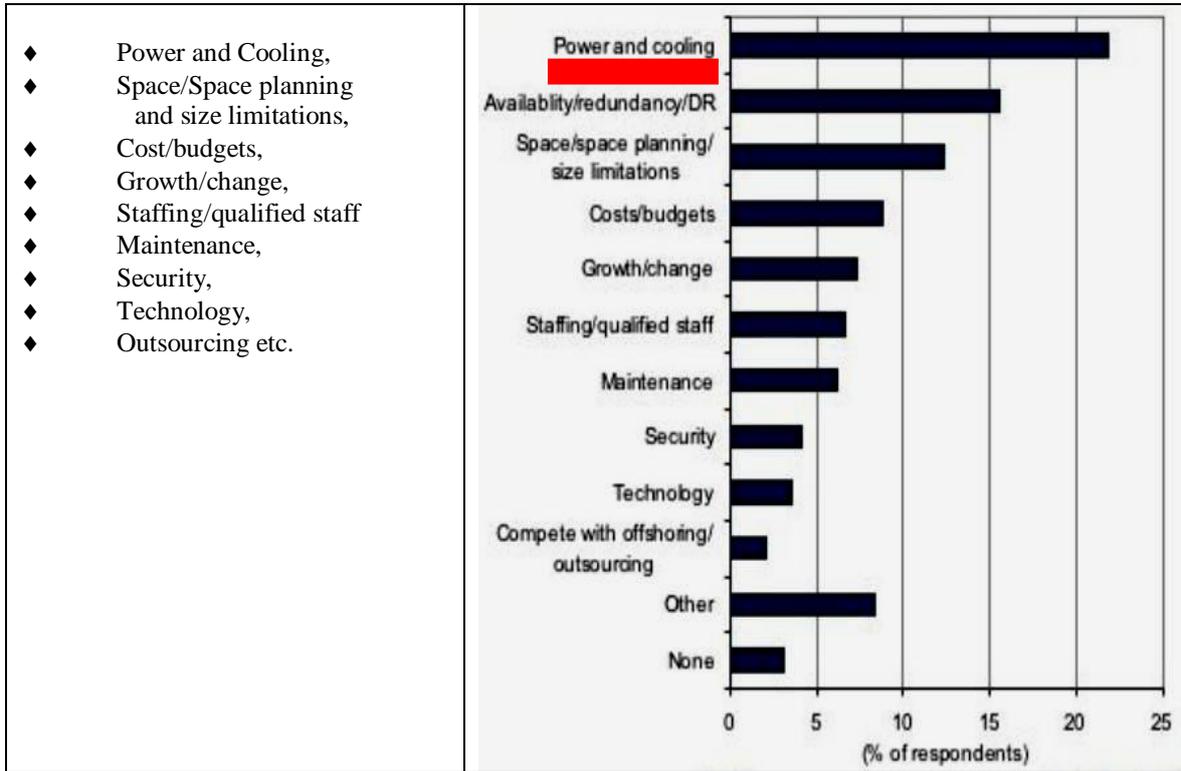


Table I. Data Center Challenges

The graph is self-explanatory and can be interpreted that Power and Cooling is the top most challenge to be faced while running, studying and improving data center efficiency.

V. WHERE GOES THE ENERGY?

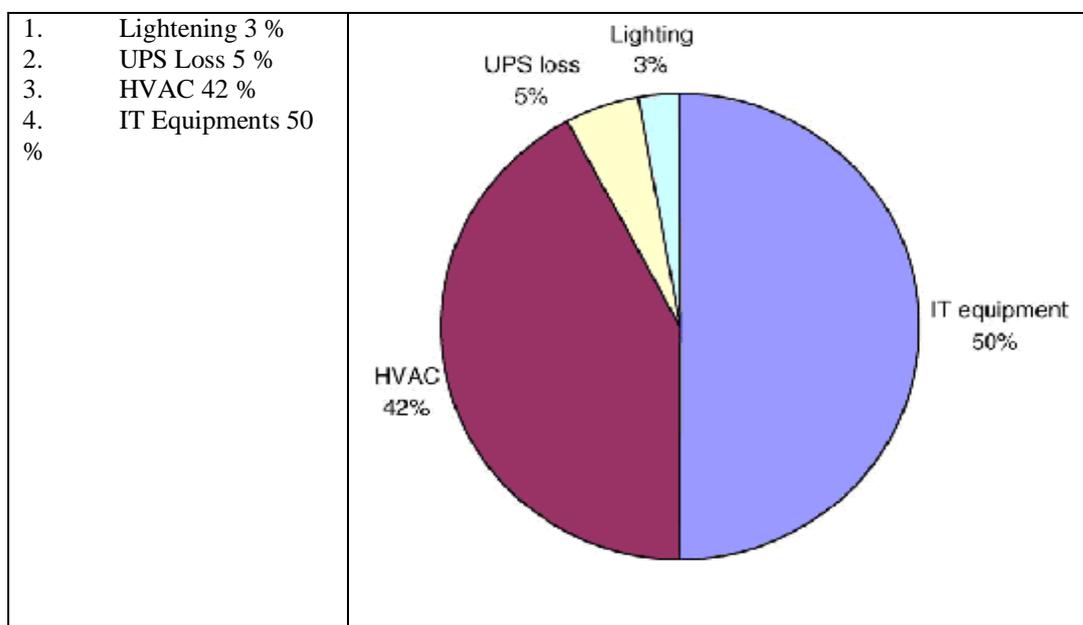


Figure II. Energy Consumption

VI. ENERGY CONSERVATION AND MANAGEMENT

5.1 HVAC LOAD CONSIDERATIONS AND CHALLENGES

HVAC loads are a high internal sensible heat load from the datacom equipment itself and a correspondingly high sensible heat ratio. The effect of the other loads (envelope, lighting, etc.) becomes proportionately more important in terms of part-load operation. The part- and low-load conditions must be well understood and equipment selected accordingly. • Existing applications floor space • Performance growth of technology based on footprint • Processing capability compared to storage capability • Change in applications over time • Asset turnover • Ventilation and Infiltration • Electrical Equipment • Lights

5.2 AIR CONDITIONING SYSTEM DESIGN AND CAPACITY

Design Criteria: Environmental requirements of datacom equipment vary depending on the type of equipment and/or manufacturer. However, a consortium of manufacturers has agreed on a set of four standardized conditions (Classes 1 to 4). A fifth classification, the Network Equipment—Building Systems (NEBS) class, is typically used in telecommunications.

Class 1	Class 2	Class 3	Class 4	NEBS
enterprise servers and storage products.	small servers, storage products, personal computers and workstations.	personal computers, workstations, laptops, and printers.	industrial controllers, or computers and handheld electronics such as PDAs.	environment are switches, transport equipment, and routers.

(NEBS) -Network Equipment—Building Systems

Table II. Datacom equipment classes

Data Processing and Electronic Office Areas

Table 1 Class 1, Class 2, and Selected NEBS Design Conditions

Condition	Classes 1 and 2		NEBS	
	Allowable Level	Recommended Level	Allowable Level	Recommended Level
Temperature control range	15 to 32°C ^{a,f} (Class 1) 10 to 35°C ^{a,f} (Class 2)	20 to 25°C ^a	5 to 40°C ^{e,f}	18 to 27°C ^d
Maximum temperature rate of change	5 K/h ^a		30 K/h ^{e,c} 96 K/h ^{e,d}	
Relative humidity control range	20 to 80%, 17°C max. dew point ^a (Class 1) 21°C max. dew point ^a (Class 2) ^a	40 to 55% ^a	5 to 85%, 28°C max. dew point ^e	Max 55%
Filtration quality	65%, min. 30% (MERV 11, min. MERV 8) ^b			Min. 85% (Min. MERV 13) ^b

Table III: Data Processing and Electronic Office Areas

5.3 TEMPERATURE MEASUREMENT IN DATACENTER

Measurement of temperature plays important role in analyzing, controlling and deciding the direction of cooling. The air temperature of the IT server room and the surrounding temperatures of the equipment were measured using a thermal infrared imaging camera and a digital thermo hygrometer.

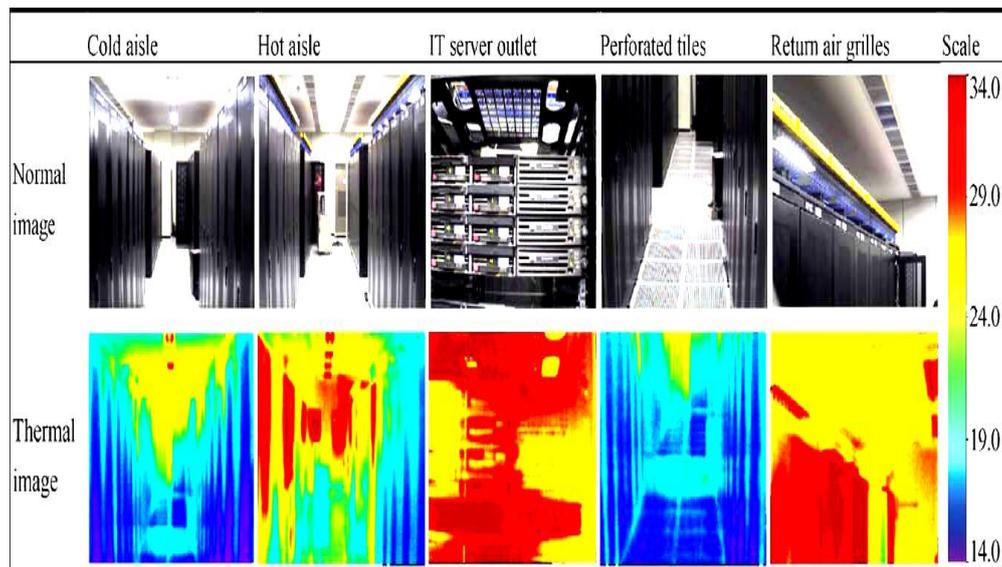


Figure III. Infrared thermography images of the IT server room.

Measurement of temperature with the help of Digital thermometers, Thermocouples, Sensors working in different temperature and environment may help to find correct temperatures.

5.4 INFRASTRUCTURE OBSERVATIONS AND MEASUREMENTS

- Cable Management • Fire Protection • Floor Layout • Overhead Ducting • Raised Floor
- Environmental Specifications • Temperature • Temperature Rate of Change Humidity • Filtration and Contamination
- Ventilation • Envelope Considerations • Pressurization. • Space Isolation. • Vapor Retarders. • Sealing • Condensation on exterior glazing. • Human Comfort • Acoustics

5.5 ANALYSIS USING LATEST TECHNOLOGIES AND BENCHMARKS

- CFD (Computational Fluid Dynamics): - Use of CFD helps to identify and automate the HAVC system controls.

• PUE / DCiE Benchmarking

PUE(Power Usage Efficiency) = Total Facility Power / IT Equipment Power

DCiE (Data Center's Infrastructure Efficiency) = IT Equipment Power / Total Power

Develop a testing schedule:

- Basic Efficiency Program: Monthly/Weekly • Intermediate Efficiency Program: Daily • Advanced Efficiency Program: Continuous (hour to hour) • Whether the calculations happen once a month or once an hour, any regular measurement is a step in the right direction.

• Best Practice Assessment

- Plan Your HVAC Efficiency Objectives • Know the HAVC power distribution components • Find your Total HVAC Power • Measuring data center HVAC performance • Recommended Data Center HVAC Performance • Measurements and Solutions • Ability to Track and Assess Equipment Availability • Ability to Assess Availability/Reliability of Critical Facilities • Meeting the Challenge • How Can The Latest Technologies Help You Improve Data Center HVAC Performance?

VII. CONCLUSION

- ◆ Coordinated and expertise efforts during early stage of designing as well as implementing (such as telecommunications, power, architectural in-outs, and supporting systems) helps to conserve and manage energy requirements for HVAC.
- ◆ For more detailed results, temperature and airflow closer to the electronic components have to be investigated using latest technology methods/procedures/equipments.
- ◆ CFD and Energy simulations tools find greater scope in research of energy conservation and management of HVAC systems.

References

1. ASHRAE. HVAC Applications, "Data Processing and Electronic Office Areas", (2007) 17.1- 17.5
2. ASHRAE. 2007. HVAC Applications, "Computer Applications", (2007) 39.1-39.16
3. H.S. Sun, S.E. Lee, *Case study of data centers' energy performance*, H.S. Sun, S.E. Lee / Energy and Buildings 38 (2006) 522–533
4. J. Fredrik Karlsson, Bahram Moshfegh, *Investigation of indoor climate and power usage in a data center*, J.F. Karlsson, B. Moshfegh / Energy and Buildings 37 (2005) 1075–1076 1083
5. Jinkyun Cho, Taesub Lim, Byungseon Sean Kim, "Measurements and predictions of the air distribution systems in high compute density (Internet) data centers", J. Cho et al. / Energy and Buildings 1108 41 (2009) 1107–1115

6. Kailash C. Karki, Suhas V. Patankar, "Airflow distribution through perforated tiles in raised-floor data centers", K.C. Karki, S.V. Patankar / Building and Environment 41 (2006) 734–744
7. Arman Shehabi a,b, Srirupa Ganguly, Lara A. Gundel, Arpad Horvath, Thomas W. Kirchstetter, Melissa M. Lunden, William Tschudi, Ashok J. Gadgil, William W Nazaroff, "Can combining economizers with improved filtration save energy and protect equipment in data centers?". Shehabi et al. / Building and Environment xxx 2 (2009) 1–9
8. Svend Frølund, Fernando Pedone, "Journal of Parallel and Distributed Computing", Volume 63, Issue 11, November 2003, Pages 1064-1081
Huirong Fu, Ming Zhang, "Online adaptive firewall allocation in internet data center" H. Fu, M. Zhang / Computer Communications 29 (2006) 1858

Web site References:

- (1) Brill, Kenneth; Pitt, Turner; and Seader, John. "Tier Classifications Define Site Infrastructure Performance," Uptime Institute, http://www.42U.com/quicklink.aspx?url=http://www.uptime.com/file_downloads/PDF/Tier_Classification.pdf
- (2) "Enterprise Server and Data Center Energy Efficiency Initiatives" <http://www.42U.com/quicklink.aspx?url=http://www.energystar.gov/index.cfm?>

Books

1. "ASHRAE Handbooks", www.ashrae.org/publications
2. "Modern Electronic Instrumentation and Measurement Technique" by A.D. Helfrick and W.D. Cooper.
3. "Experimental Methods For Engineers", McGraw Hill International Edition, Seventh Edition by J.P.Holman.
4. "Energy Auditing and Energy Management" by Abby Y.P.

Analysis & Optimization of Crankshaft Using Fem

R. J. Deshbhratar¹, Y. R. Suple²

*IVth Sem M.Tech Student, Kavikulguru Institute Of Technical And Science, Ramtek
Asst. Prof. Kavikulguru Institute Of Technical & Science, Ramtek*

Abstract: *The modal analysis of a 4-cylinder crankshaft is discussed using finite element method in this paper. The analysis is done on two different materials which are based on their composition. Three-dimension models of diesel engine crankshaft was created using Pro/ENGINEER software. The finite element analysis (FEM) software ANSYS was used to analyse the vibration modal of the crankshaft. The maximum stress point and dangerous areas are found by the deformation analysis of crankshaft. The relationship between the frequency and the vibration modal is explained by the modal analysis of crankshaft. The results would provide a valuable theoretical foundation for the optimization and improvement of engine design.*

Keywords: *crankshaft, finite element analysis; optimization; Thermal analysis*

I. INTRODUCTION

Crankshaft is one of the most important moving parts in internal combustion engine. It must be strong enough to take the downward force of the power stroked without excessive bending. So the reliability and life of internal combustion engine depend on the strength of the crankshaft largely. And as the engine runs, the power impulses hit the crankshaft in one place and then another. The torsional vibration appears when a power impulse hits a crankpin toward the front of the engine and the power stroke ends. If not controlled, it can break the crankshaft. Strength calculation of crankshaft becomes a key factor to ensure the life of engine. Beam and space frame model were used to calculate the stress of crankshaft usually in the past. But the number of node is limited in these models. With the development of computer, more and more design of crankshaft has been utilized finite element method (FME) to calculate the stress of crankshaft. The application of numerical simulation for the designing crankshaft helped engineers to efficiently improve the process development avoiding the cost and limitations of compiling a database of

real world parts. Finite element analysis allows an inexpensive study of arbitrary combinations of input parameters including design parameters and process conditions to be investigated. Crankshaft is a complicated continuous structure. The vibration performance of crankshaft has important effect to engine. The calculation

II. 3-D entity model of crankshaft

The structure of the crankshaft has more small fillets and fine oil hole. Considering these factors in establishment process, finite element mesh of crankshaft becomes very densely, the number of node equation increase greatly. These factors would extend the solution time, make the unit shape unsatisfactory and amplify the accumulative error. This would lower the simulation accuracy. Hence,

of crankshaft vibration performance is difficult because of the complexity of crankshaft structure, the difficult determinacy of boundary condition. Dynamic matrix method and dynamic substructural method combined with FME were used to calculate the vibration of crankshaft. The method of three-dimensional finite element was carried to analyse dynamical characteristic of diesel crankshaft.

In the paper, 3-D finite element analysis are carried out on the modal analysis of crankshaft and the thermal analysis of crankshaft, And the FME software ANSYS was used to simulate the modal analysis of the crankshaft. The results of natural frequencies and mode shape were obtained. And deformation distributions of crankpin were obtained by using ANSYS software. The results are regarded as a theory basis to optimize the design of crankshaft and thermal analysis of crankshaft.

CRANKSHAFT MODELLING

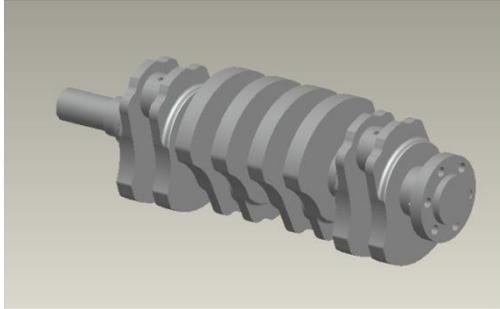
In the present research, diesel engine crankshaft was studied. The crankshaft has four crankthrows, three rod journals and two main journals, and the mainly dimension parameters are considered while preparing model in Pro/E. According to complicated structure of crankshaft, the integral crankshaft should be applied when performing finite element model analysis. In view of the structure shape characteristics of crankshaft, reducing the solution time and enhancing the simulation accuracy, the crankshaft model was simplified in establishment process. The premise of simplification is that it does not influence the dynamics characteristics of crankshaft. The crankshaft model was created by Pro/ENGINEER software.

model was meshed by 8 unit solid45. the meshing accuracy is 3 grade. After automatic meshing, in order to enhance the simulation accuracy and avoid generating the macrocephalic mesh, the crankshaft dangerous areas were further tessellated. The 3-D crankshaft finite element grid model is shown in Figure. The two materials different materials for crankshaft is used. The physical parameters used in the crankshaft simulation were list in Tables 01 and Table 02.

Then the model was imported to the ANSYS software. According to the structure of crankshaft, the crankshaft diameter less than 12mm were ignored. The model of four-cylinder crankshaft is shown in Fig.

in the paper, the real crankshaft was represented by a simplified model. In this simplified model, the chamfers which radius less than 5mm and the oil holes which

Meshing of 3-D Entity



FINITE ELEMENT METHOD

The finite element method is numerical analysis technique for obtaining approximate solutions to a wide variety of engineering problems. Because of its diversity and flexibility as an analysis tool, it is receiving much attention in engineering schools and industries. In more and more engineering situations today, we find that it is necessary to obtain approximate solutions to problems rather than exact closed form solution. It is not possible to obtain analytical mathematical solutions for many engineering problems. An analytical solutions is a mathematical expression that gives the values of the desired unknown quantity at any location in the body, as consequence it is valid for infinite number of location in the body. For problems involving complex material properties and boundary conditions, the engineer resorts to numerical methods that provide approximate, but acceptable solutions.

The finite element method has become a powerful tool for the numerical solutions of a wide range of engineering problems. It has developed simultaneously with the increasing use of the high-speed electronic digital computers and with the growing emphasis on numerical methods for engineering analysis. This method started as a generalization of the structural idea to some problems of elastic continuum problem, started in terms of different equations or as an extrinum problem

The fundamental areas that have to be learned for working capability of finite element method include:

- Matrix algebra.
- Solid mechanics.
- Variation methods.
- Computer skills.

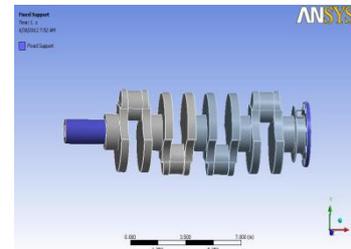
Matrix techniques are definitely most efficient and systematic way to handle algebra of finite element method. Basically matrix algebra provides a scheme by which a large number of equations can be stored and manipulated. Since vast majority of literature on the finite element method treats problems in structural and continuum mechanics, including soil and rock mechanics, the know-ledge of these fields became necessary. It is useful to consider the finite element procedure basically as a Variation approach. This conception has contributed significantly to the convenience of formulating the method and to its generality.

The term "finite element" distinguishes the technique from the use of infinitesimal "differential elements" used in calculus, differential equations. The method is also distinguished from finite difference equations, for which

although the steps in to which space is divided into finite elements are finite in size; there is a little freedom in the shapes that the discrete steps can take. F.E.A is a way to deal with structures that are more complex than dealt with analytically using the partial differential equations. F.E.A deals with complex boundaries better than finite difference equations and gives answers to the „real world“ structural problems.

BOUNDARY CONDITION

The crankshaft bears the constraints of main journals and longitudinal thrust bearing. Because of the effect of load, crankshaft main journals appear bend deformation between the lower main-bearing half and upper main-bearing half. And the longitudinal thrust bearing can prevent effectively the crankshaft axial movement and ensure the piston-and-connecting-rod assembly normally work. Five surface radial symmetry constrains were exerted on the five main journals surface respectively, axial displacement constrains were exerted on the two end face of crankshaft. Then the modal analysis was carried out using the ANSYS software.



CONSTANT PROPERTIES OF BOTH MATERIAL

Table No.01

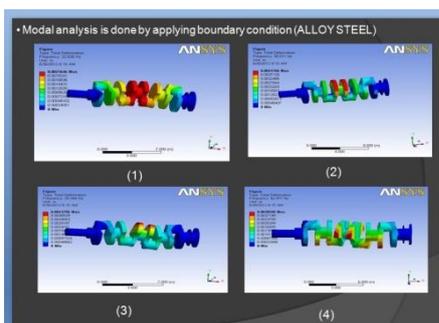
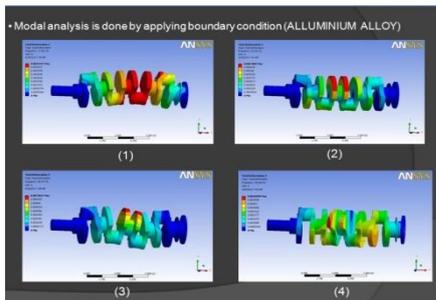
Aluminum Alloy Constant Properties	
Name	Value
Compressive Yield Strength	2.8×10^8 Pa
Density	2,770.0 kg/m ³
Poisson's Ratio	0.33
Tensile Yield Strength	2.8×10^8 Pa
Tensile Ultimate Strength	3.1×10^8 Pa
Young's Modulus	7.1×10^{10} Pa
Thermal Expansion	2.3×10^{-5} 1/°C
Specific Heat	875.0 J/kg·°C
Relative Permeability	10,000
Resistivity	5.7×10^{-8} Ohm·m

Table No. 02

Alloy Steel properties	constant
Name	Value
Compressive Yield Strength	2.5×10^8 Pa
Density	7,850.0 kg/m ³
Poisson's Ratio	0.3
Tensile Yield Strength	2.5×10^8 Pa
Tensile Ultimate Strength	4.6×10^8 Pa
Young's Modulus	2.0×10^{11} Pa
Thermal Expansion	1.2×10^{-5} 1/°C
Specific Heat	434.0 J/kg.°C
Relative Permeability	10,000.0
Resistivity	1.7×10^{-7} Ohm·m

IV. RESULT AND DISCUSSION

The first four modal frequencies for both the material is calculated from ansys



V. CONCLUSIONS

In this paper, the crankshaft model is created by Pro/ENGINEER software. Then the model created by pro/Engineer was imported to ANSYS software. The maximum deformation appears at the centre of crankshaft surface. The maximum stress appears at the fillets between the crankshaft journal and crank cheeks, and near the central point. journal. The edge of main journal is high stress area. The crankshaft deformation was mainly bending deformation under the lower frequency. And the maximum deformation was located at the link between main bearing journal and crankpin and crank cheeks. So this area prones to appear the bending fatigue crack. Base on the results, we can forecast the possibility of mutual interference between the crankshaft and other parts. The resonance vibration of system can be avoided effectively by Appropriate structure design. The results provide a theoretical basis to optimize the design and fatigue life calculation

ACKNOWLEDGEMENTS

The support extended by the guide (Shri Y.R. Suple) and college authorities is highly appreciated and acknowledged with due respect.

REFERENCES

- [1] Amit Solanki, Ketan Tamboli, M.J.Zinjuwadia, 'Crankshaft Design And Optimization- A Review' ,National Conference On Recent Trends In Engineering & Technology.
- [2] C.M.Balamurugan, R.Krishnaraj, Dr.M.Sakthivel, K.Kanthavel, Deepan Marudachalam , 'Computer Aided Modelling And Optimization Of Crankshaft' International Journal Of Scientific & Engineering Research Volume 2, Issue 8, August- 2011 Issn 2229-5518.
- [3] Vladimir V. Riabov , 'Computational And Analytical Methods In The Engine Crankshaft Design' ,Rivier College, Nashua, New Hampshire 03060, Usa.
- [4] Jian Meng, Yongqi Liu, Ruixiang Liu Zibo, Shandong, China 'Finite Element Analysis Of 4-Cylinder Diesel Crankshaft'I.J. Image, Graphics And Signal Processing, 2011, 5, 22-29 Published Online August 2011 In Mecs
- [5] ZHOU Xun, YU Xiao-li , 'Reliability Analysis Of Diesel Engine Crankshaft'Based On 2d Stress Strength Interference Model, School Of Mechanical And Energy Engineering, Zhejiang University, Hangzhou 310027, China.
- [6] Jenson (1970), ' Dynamic Load Determination And Analysis'
- [7] Mehrdad Zoroufi and Ali Fatemi, 'A Literature Review on Durability Evaluation of Crankshafts Including Comparisons of Competing Manufacturing Processes and Cost Analysis', 26th Forging Industry Technical Conference, Chicago, IL, November 2005.

3D Reconstruction of Human Retina from Fundus Image – A Survey

Jini Cheriyan¹, Hema P Menon², Dr. K. A. Narayanankutty³

1. Asst. Professor, Dept of ECE

2. Asst. Professor (SG), Dept of CSE, 3. Professor, Dept of ECE

ABSTRACT: Imaging techniques for the eye include fundus camera, OCT etc of which the most widely used and economical one is the use of fundus camera. Images obtained from this camera are 2D in nature. Analysis of 2D data requires a lot of expertise. This makes it necessary for a 3D reconstruction of the 2D fundus image which would help doctors in their analysis and treatment plan. 3D retinal image would also be helpful in explaining to patients, the progression of disease. In this paper a survey on various methods for 3D reconstruction of retinal fundus image has been discussed.

I. INTRODUCTION

A visual image is rich in information. 3D visualization technique is suitable for the display of complex structures. 2D visualization of images especially in medical images doesn't give significant information about the object and their properties, which are very useful in diagnosis and treatment. The 2D representation lacks showing different number of views, information loss due to lack of giving depth information of the object., doesn't give much information about the structure of the object, and lack of giving realistic effect. These drawbacks can be overcome by developing 3D models, that have an additional z direction which gives the depth information of the object which is very useful in medical image visualization.

The doctors can visualize a portion or area suspected to be infected by the disease in 3D view to examine the spread of the disease. Accurate diagnosis often requires a 3D image of retina. The analysis of the 3D shape of retinal fundus is required for identifying lesions and estimating the intensity of lesion. The produced 3D model of the eye can be used for studying the anatomy of the eye diagnosis of retinal disease, treatment planning and for educational purpose.

II. LITERATURE SURVEY

2.1 IMAGING TECHNIQUES FOR EYE

2.1.1 HYPERSPECTRAL FUNDUS IMAGES

A fundus camera is used to photograph the interior surface of the eye. Fundus cameras are used by optometrists and ophthalmologists for monitoring progression of a disease and diagnosis of a disease. Fundus image of retina provides physio-pathological information and is used to detect exudates and haemorrhages. The observation of fundus is used for diagnosis of eye diseases and also for checking the whole body conditions. Figure 2.1.1 shows the fundus images of human retina.

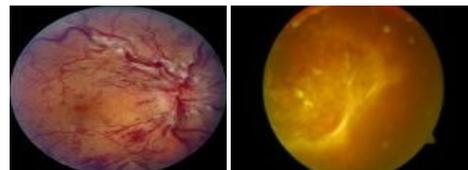


Figure 2.1.1: Fundus images of human pathological retina

2.1.2 OPTICAL COHERENCE TOMOGRAPHY

Optical Coherence Tomography, or 'OCT', is a 3D imaging technique for obtaining sub-surface images of opaque materials at a low-power microscopic resolution. OCT has become a leading imaging method in displaying the structures of the eye. Figure 2.1.2 shows the OCT images of retina. OCT is very useful for the medical community, because it provides tissue level information at higher resolution than other imaging modalities such as MRI or ultrasound. OCT image shows the quantity and configuration of macular holes, intra-retinal layers intra-retinal cysts and photoreceptor layer. Main advantage of using OCT is that it has no ionizing radiation.

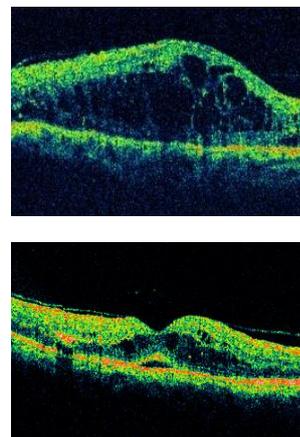


Figure 2.1.2 : OCT images of human retina

2.1.3 INTRAVENOUS FLUORESCENCE ANGIOGRAM (IVFA)

Intravenous **fluorescein angiography** (IVFA) is another imaging technique commonly used by ophthalmologists for evaluating the internal blood circulation of eye. IVFA is helpful in detecting leaks and abnormalities. IVFA are commonly used in certain cases of diabetic retinopathy and age-related macular degeneration. The photographic record of IVFA is called as fluorescein angiogram.

2.2 3D MODELLING TECHNIQUES

2.2.1 Y-FEATURE BASED METHOD

To obtain the 3-D shape of retina, practitioners generally use the specialized devices such as the scanning laser ophthalmoscope and the optical coherence tomography (OCT) system which provides a cross-sectional image of retinal fundus. The use of OCT is not widespread because it require costly equipment .Therefore the alternative method consists of inferring the 3-D shape of the retina using fluorescein images of the retina.

Existing literature reveals that atleast a pair of retinal images are required for 3D reconstruction of the retinal image. In Choe et al [1] paper, they have proposed a method which is a computer vision based approach that extracts the location of vessels bifurcation . For extracting vessel bifurcation plane and parallax method is used. Then estimates the fundamental matrix for nearly planar surface, the retinal fundus. The use of mutual information is used for the estimation of the dense disparity maps, where the matched Y-features are used for estimating the bounds of the range space disparity. Y-feature is the most commonly used feature since it is easy to detect and well distributed in fluorescein images. The method consists of three steps: First, accurate positions of bifurcation of vessels are extracted and matched across images using an articulated Y-feature model. Second, using the matched Y-features, a plane-and-parallax approach is considered for estimating the epipolar geometry, and the search space on the scan line for stereo matching is estimated. Subsequently, a dense disparity map is estimated by matching point using a mutual information method. The 3-D shape of lesions, a fovea, and an optic disc are also accurately estimated.

2.2.1.1 Y-FEATURE EXTRACTION

Y-feature are the regions where three vessels converge. seed positions of y-feature are located by using PCA based analysis for every pixel.The location with third largest eigen value are taken as the seed points. Y-feature has 8 DOF . And fit the model using gradient descent method by minimizing the energy.

$$f(x) = (-1)^m \frac{1}{2} \sum_{i=1}^3 \int_{-w_i}^{w_i} \int_0^l [I(x_i, y_i)]^2 dl dw$$

$$- \frac{1}{2} \sum_{i=1}^3 \sum_{w \in (-w_i, w_i)} \int_0^l [G(x_i, y_i)]^2 dl \text{-----(1)}$$

Where i is the index of the arm considered, I(xi, yi)is the intensity ,G (xi,yi) is the gradient value . m=0 indicates dark vessel and m=1 indicates bright vessel.

$g_j(\Theta)$ is the angle between each arm, bl and bu are the range values for angles and wl and wu are range values for widths.

$$B(\Theta, w) = \sum_{j=1}^3 \left(\frac{1}{(g_j(\Theta) - bl)(bu - g_j(\Theta))} + \frac{1}{(w_j - wl)(wu - w_j)} \right) \text{----(2)}$$

$$E(x) = f(x) + \lambda B(\Theta, w) \text{----(3)}$$

λ

is the trade off value. The energy function E is minimized using gradient method. Next step is to match the Y-features of source and target imge. For that consider two windows Za and Zb for source and target respectively

$$MI(Zz, Zb) = H(Za) + H(Zb) - H(Za, Zb) \text{----(4)}$$

where H(Z) is the Shannon entropy of the window. After finding the mutual information homography and fundamental matrix are to be estimated.

2.2.1.2 ESTIMATION OF EPIPOLAR GEOMETRY

RANSAC method is used to detect the inliers among the matched pairs. And select the best four pairs to estimate the homography in order to minimize the geometric error. In order to find the fundamental matrix plane and parallax based on RANSAC is used.

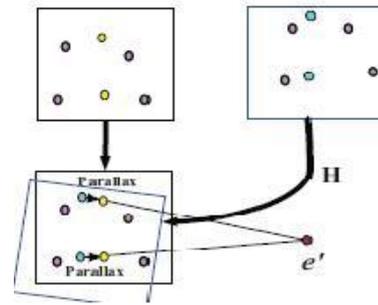


Figure2.2.1.2 estimation of fundamental matrix using plane and parallax method

$$F = [e']H$$

e' where is the epipole. After calculating the fundamental matrix stereo images are rectified for the estimation of depth map.

2.2.1.3 STEREO MATCHING

Since retinal fundus has small depth variations ,disparity range is also very narrow. Search space is commonly taken as the maximum disparity gradient value. But this cannot be taken in the case of retinal images. We know the disparity at some key features so it is possible to find the disparity at other points since retinal surface is a smooth and curved surface. In order to estimate the depth map mutual information for matching points are used.

$$disparity = \max MI(d) = \max MI(Za, Zb, d)$$

Zb,d is the window with d pixel distance from Zb. The window pairs that has the maximum mutual information is used to determine the disparity of each pixel location. In order to avoid the staircasing effect in the disparity map of retinal images subpixel resolution is considered.

$$disparity = \max MI(d) = \max MI(Za, Zb, d)$$

Where
 a=MI(disparity-1),
 b=MI(disparity),
 c=MI(disparity+1)

Each disparity is calculated individually by using mutual information mathching. By using this method different parts of retinal fundus such as fovea,optic disc lesions are accurately estimated.

2.2.2 RECONSTRUCTING 3D USING SEQUENCE OF ANGIOGRAMS

For reconstructing the 3D surface information of the human retinal fundus using a sequence of fluorescein angiograms is presented in [2] by F Laliberte .These angiograms are taken with an uncalibrated camera. In this work the camera is still and assuming that the natural head and eye movement is high to create different views as that of stereo. The results obtained shows the distribution of fluorescein with in the retinal fundus. This 3D fluorescein distribution obtained gives information about circulation and staining. This method is mainly concentrated on the macular region of retina rather than a full image. Data used include six sequence of angiograms of 3 patients having macular degeneration. The output obtained by using this method is a dense disparity map, dense disparity map is an image in which each pixel is the difference in distance between the corresponding part of two original images. The overall reconstruction procedure described in this paper include i) detection of control points and find the matching points between each image and the reference image. ii) calculate the disparity iii) rectification. Here the control points used are bifurcation points of the retinal blood vessels. For calculating the fundamental matrix eight point algorithm is used. The information obtained through the dense disparity map is same as the visual assessment done by the ophthalmologist.

2. 2. 3 3D RECONSTRUCTION BY REGISTRATION

Another method for 3D reconstruction which is based on registration is discussed in paper [3] . Here registration is based on both the area and feature in order to obtain an optimal solution and a translational model is estimated through binary mutual information instead of gray level mutual information. In this work a weak perspective camera is assumed because each retinal image has small depth variation. Then the geometric constraint of an eyeball is taken in to account to generate denser points for surface. Since the eyeball is approximated as a sphere a point based sphere fitting method is used. The author proposed a new algorithm for extracting blood vessels, it include matched filtering, local entropy thresholding ,length filtering and bifurcation and this algorithm does not involve any human interaction and less computational complexity. Prior to 3D reconstruction affine bundle adjustment is done to produce jointly optimal structure and viewing parameters and recovered a retinal euclidean surface

2.2.4 3D FUNDUS RECONSTRUCTION AND DISPLAY FROM MULTIPLE FUNDUS IMAGES

Koicchio et al in paper [4] proposed a method to display 3D fundus from a set of multiple partial images of fundus. The different views are obtained by shifting the fundus camera. This method used the concept that the fundus has a spherical shape and the image of eye lens results in a quadratic surface. This method calculates all the optical parameters. A combination of eye lens and the enlarging contact lens are modelled and identified the optical parameters of modelled lens. Then the spherical surface of the fundus is mapped on to a quadratic surface as its real image. Then extracts the feature points from fundus image and find correspondences and the corresponding pairs are registered on a quadratic surface. Parallely the viewing

direction of camera to take individual images is identified to produce best corresponding pairs. The fundus pattern are back projected from the multiple images to the reconstructed sphere.

III. FINDINGS AND OBSERVATIONS

Existing 3D reconstruction methods require multiple fundus images , taken from different orientations. Since acquisition of image needs the patients concent and sometimes it may be difficult, and processing of images consumes a longer time. Reconstruction of single image minimizes acquisition and processing time. New methods have to be identified for reconstruction from single image. The authors are continuing their research in this field.

IV. CONCLUSION

An extensive survey on the various methods for 3D reconstruction of retinal fundus image has been discussed in this paper. These informations would be useful for researchers working in the field of 3D image reconstruction.

ACKNOWLEDGEMENTS

The authors would like to thank Dr.Gopal S pillai, Department of Ophthalmology and Dr. Kumar Menon, Incharge, Center for Digital Health , AIMS, Cochin, India, for their support in conducting this survey

REFERENCES

- [1] Tae Eun Choe,Issac Cohen,Gerard Medioni,"3Dshape reconstruction of retinal fundus" *IEEE computer society conference on computer vision and pattern recognition*,2006
- [2] F. Laliberté, L. Gagnon, Y. Shengb R&D Department, CRIM, 550 Sherbrooke West, Montreal, QC, Canada, H3A 1B92 Physics Department, Laval University, Quebec, QC, Canada, G1K 7P4, "Three-dimensional visualization of human fundus from a sequence of angiograms"
- [3] Thitiporn Chanwimaluang , "Advanced retinal imaging: feature extraction, 2-D registration, and 3D reconstruction"
- [4] Koichiro DEGUCHI ,Z Junko NOAMIZ, Hidekata HONTANIZ , " 3D Fundus Pattern Reconstruction and Display from Multiple Fundus Images", *IEEE*; 2000
- [5] M. Wunstel, H. Schumann, "Automatic 3D-reconstruction of the ocular fundus from stereo images", *Proc. Computer Assisted Radiology and Surgery*, pp. 456-60 ,2002
- [6] Kitara Kadhim Al-Shayeh and Muzhir Shaban Al-Ani, "Efficient 3D Object Visualization via 2D Images", *IJCSNS International Journal of Computer Science and Network Security*; VOL.9 No.11, November 2009
- [7] K. Deguchi, D. Kawamata, K. Mizutani, H. Hontani, K. Wakabayashi," 3D fundus shape reconstruction and display from stereo fundus images", *IEICE*

- Transactions on Information and Systems*, E83-D, pp. 1408-14, 2000
- [8] Kolar R, Kubecka ,Jan J ;,"Registration and fusion of the autofluorescent and infrared retinal images". *Image and vision computing*; 2008
- [9] Radim Kolar,Viktor Sikula,michal Base; " Retinal image registration using correlation" *BIOSIGNAL*; 2010
- [11] France Laliberté, Langis Gagnon" Registration and Fusion of Retinal Images—An Evaluation Study" *IEEE TRANSACTIONS ON MEDICAL IMAGING*, VOL. 22, NO. 5, MAY 2003
- [12] A. Bartoli, P. Sturm "Nolinear Estimation of the Fundamental Matrix with Minimal Parameters," *IEEETrans. on PAMI*, Vol. 25, No. 3, March 2004
- [13] R. Hartley, "In Defence of the 8 point Algorithm", *ICCV1995*, pp.1064-1070, June 1995
- [14] H. Hirschmüller, "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information," *CVPR 2005*, pp. 807 – 814, 2005
- [15] W.M. Wells III, P. Viola, H. Atsumi S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, March 1996
- [16] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? *ECCV (1)*, pages 578– 591, 2006
- [17] L. Zhang, G. Dugas-Phocion, J. Samson, and S. Seitz. Single view modeling of free-form scenes. *CVPR01*, pages 1:990–997, 2001.
- [18] Yuji Iwahori, Takashi Nakagawa, Robert J. Woodham , Shinji Fukui, Haruki Kawanaka "Shape from Self-Calibration and Fast Marching Method" *IEEE*,2008
- [19] F. Zana and J. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and hough transform," *IEEE Trans. Biomedical Engineering*, vol. 18, pp. 419-428, May 1999.
- [20] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, "A feature-based,robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 347-364,March 2002.
- [21] D. Wu, M. Zhang, J. C. Liu, and W. Bauman, "On the adaptive detection of blood vessels in retinal images," *IEEE Trans. Biomedical Engineering*, vol. 53,pp. 341-343, February 2006.
- [22] J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proc. Nat. Acad. Sci.*, 93(1):5– 28, 1998.

Assessment of the extent of using cloud computing technology in economic applications

Dr. Maha Mahmoud Talaat mustafa¹

*(information systems & computer Department, Sadat Academy for management sciences, Egypt)

Abstract: The objective of this research is to assess the factors that will affect the economic status of the organizations as a result of using cloud computing systems. Among these factors, which will be examined in this study, the relative economic benefit expected through the application of cloud computing, which relies primarily on the degree of complexity or simplicity in the work, the compatibility between the components of information system, top management support, the expected growth in the volume of business, technological preparations, and finally competitive advantages achieved.

The research also aims to help organizations to make financial and economic analysis, to identify several economic characteristics of their work, as well as to identify the resources required to shift to cloud computing.

IT portfolio Model has been developed according to cloud computing system

Also cost-benefit analysis has been used in order to perform the economic assessment, taking into account the impact of the intangible resources and intangible benefits at the expense of investment in cloud computing. The cost-benefit analysis has provided a broader vision for the work outcomes in the cloud computing systems.

Keywords: Benefits of cloud computing, Cloud computing, Cost of cloud computing, IT Portfolio,

I. INTRODUCTION

What is cloud computing?

Cloud computing system, is a global system that enables the user to access the network and directly use its information from anywhere, anytime around the world, by a special login name and password.

Economic advantages of solutions based on cloud computing

1 – Take advantage of the possibilities in cyber space

Of the advantages of cloud computing, as a user navigates to the international network to the wide world of cyber space, and the creation of a unified center for programs; this system reduces power consumption, keeps the programs from the loss, in addition to the continuity of use.

2-savings in both software licensing and hardware, as well as to dispense with the full use of the structure of a computer key, and thus saving in the spatial area and electricity, and also helps in reducing costs and risks associated with the project update, through a change in the portfolio of information technology

3- Providing another alternative for the construction of the infrastructure for the portfolio of information technology at less cost .

Where the service provider of information technology can help the organization to develop the infrastructure of Information technology, especially computers and software, by allowing organizations to focus on business.

4-Optimal exploitation of the capacity of servers

The capacity of the server can be enhanced in business applications, therefore avoiding the need to invest in expensive hardware and software, which allows the organization to focus its attention on core business .

5-Providing more flexibility to work remotely

Cloud computing enables users to work from anywhere as long as they are connected to the internet, which means provision of greater flexibility in work schedule, and reducing stress.

Organizations will be more flexible such that they can operate more efficiently, by making it easier to rapid response to changing market needs, and internal applications, and access to new technologies, as well as easier deployment, lower maintenance, and confidence in the applied solutions, which are carried out under the supervision of a service provider of information technology. In general we can summarize the advantages and disadvantages of cloud computing as follows:

II. ADVANTAGES OF CLOUD COMPUTING

The cloud computing system provides the following benefits :

Cloud computing services enabled office users to access resources from any computer, (Gillam.2010) which provides many of the multiple benefits :

1. Workers can continue their work in the event that access to their work is hindered due to the presence of obstacles.
2. Companies can make use of staff across a wider geographical area, especially if salaries are lower as in the distant localities.
3. To maintain business operations in case of interruptions as of electricity or any other adversity in the office .
4. Providing better facilities to address the problems and interruptions of electric power networks .
5. Data recovery from data centers, as in case of disasters.
6. Replication of infinite data.
7. Application of enhanced support systems, by deleting redundant data.
8. Enables electronic archiving.
9. Application of governance systems
10. Risk management
11. Regulatory compliance on IT solutions

12. Protects data from loss .
13. Leads to lower operating costs.
14. Lower costs through sharing resources

Disadvantages of Cloud Computing.

1. Ambiguity of security features in these systems
2. Levels of protection for data and information is not specific
3. Failure to maintain the security aspects
4. The need to conduct security assessments.
5. The need to assess the legal aspects in the areas of legislative compliance, audit and follow-up .
6. Vulnerability of cloud computing systems, for many cyber-attacks .

What is the virtual storage technology?

It is the ability to run multiple operating systems on a single normal system and to participate in the sources of the hard drives where data is stored.

According to the development of technologies in computer networks, which provided a link between computers and storage devices, the storage technology changed tremendously and, at the same time, the viability of new storage technology emphasizes the use of fast networked storage, making it easier to access data from a distance, and to simplify the management in a flexible manner.

Economic savings resulting from cloud computing

1. Reduction in the cost of maintenance
2. Reduction in wages of technical Support staff
3. Reduce energy bills
4. Reduce the cost of computers
5. Reduce the cost of software licenses
6. There are no additional costs to computer upgrades.

Economics of cloud computing applications

For information Technology (IT), the cost of management and energy are significantly low, due to the smaller sizes of modern computers.

In addition, the cost of maintenance for computers represents a larger amount in the information centers for electronic archiving of the contents of cultural heritage. The majority of cost is the salaries of managers who are constantly looking for ways to save power, and to increase flexibility, to ensure reductions in the costs incurred, and increasing revenues.

Since the planning and the capacity of the information centers of the traditional institutions must be made to withstand peaks of work periods; it is not appropriate to use a computer server of low capacity where the business volumes are relatively high, even in the rest of the year a large amount of less data needs to be stored. According to the follow-up work in the information centers industry, in the past few years ago, there is a tendency to central management of data, and virtual storage that may be in most of the time of the year.

The servers has allowed the spread of the fastest computers, and specialized servers lead to a higher density of the server device without increasing the size of the data center, the number of employees, or even more power consumption.

In any case, these alternatives still require significant investments and commitments of long-term technology, and there has been increasing attention paid to the use of those alternatives, including speed of deployment, and lower maintenance requirements as less as possible. The use of cloud computing systems can meet these many requirements needed by banks and information centers.

The potential economic cost associated with the use of cloud computing services

The potential economic cost associated with the use of cloud computing services includes:

The costs of securing data

The costs of securing the transfer of data, especially sensitive data, and the cost of processing or storing data remotely. The cost resulting from limitations in the development possibilities that are currently available as a result of the developments to the service providers.

Difficulties in transfer of the proprietary of data, and programs affecting the functioning of the organization from a computer server to another.

Integrating services of cloud computing with systems of work organization; the need to modify the cost of business applications, according to the organization of business applications in the cloud computing environment

The cost of transporting data from one place to another.

Economics of Cloud Computing

Client/server computing system Is a quantum leap of the client / server mainframe to a system of cloud computing and this point is similar, in importance and amount, to a shift from the mainframe hardware to the client / server (Harms And Yamartina ,2010).

Cloud computing environment provides basic information through the provision of large data centers that can take significant advantage as a result of the economics of the wide range of applications in three directions:

1- Savings from the supply side: Large data centers with lower costs for each server machine.

2- Savings from the demand side: Evaluation of the demand for cloud computing allows flexibility to take more advantage of the change in general rates.

3- Effectiveness of the plurality of leasing: When the user change to the applied multi-lease system, increasing the number of leases, means reducing the cost of management and application of servers proportionally.

Reduction in the cost:

Costs reduction in cloud computing is due to the reduction in the price of:

1- Electrical energy: where the electricity power is of the main elements that represent the cost in using computers, may represent up to 15% to 20% of the costs of the total operating costs, nevertheless, the efficiency of energy use tends to be lower for large facilities than for smaller facilities. Anyway, the cost can be minimized by locating

data in the data centers, which is located in areas where electrical energy is economical. In addition this can be achieved through contract agreements for electricity at the possible lowest cost.

2- Labor costs: Costs savings of cloud computing labor could be achieved by automating many repetitive administrative tasks. The administrations of larger facilities can reduce costs more than smaller facilities, where system administrator can coordinate to serve a larger number of workers through data computing, which gives him the opportunity to focus on other tasks and more activities that add value to the product and improve the skills of workers.

3- The cost of insurance and ensuring the validity of the cyber product: For the cloud computing, there is general need for some degree of security to ensure the confidentiality of data or information stored on the servers, in other words, an increase in the costs; nevertheless, if we have achieved savings through cloud computing, the savings attained through securing the data, must not exceed those savings

Savings from the demand side

The total cost of information technology IT is not measured only to the cost of capacity (John W. Rittinghouse, 2009) , but also to the degree by which the capacity is utilized effectively; in this case we need to assess the amount that can be circulated through demand that we can achieve, the cost of resources actually utilized, such as the storage networks, and others.

III. THE COST OF APPLICATIONS FOR EACH WORK LOAD

In data centers which do not contain all applications in one place, savings in the cost of the work load for each application work can be realized through the server device, it does not matter here the number of servers used according to each classification of downloads, but what matters is to benefit the most from servers according to the different applications of the imaginary work, which can be through the many applications by working on the servers or even a single server, and by specific operating system,

by working imaginative work on a less number of servers with the same amount of download analysis. In fact work downloads changed over time, it may be the demand for work downloads is high in some period and less in other periods, in this case, the distribution of downloads of various work must be throughout the day or over different periods, so they can make full use of those servers at all times.

Change in the work downloads in banks.

Work downloads can be changed over different periods based on several causes, including:

Changes throughout the day vary according to the following reasons:

1- Random: models of end-user to enter their account contain some degrees of randomness, for example, individuals can access their email at different times of the

day, therefore they must take into account that many people may enter on the device server at the same time.

2 - Time of day: There is a daylong cycle from eight thirty in the morning to five o'clock in the afternoon, where more pressure is from eleven am to three o'clock in the afternoon.

Therefore the working capacity should take into account the pressure on the servers during those periods.

3- Changes in the work downloads according to seasonal changes : There are many changes in the work downloads, which vary from one industry or application to another, and this difference is the one which governs the movement in industries, for example, there are industries where the pressure of work is during holidays and seasons of purchase, while for example in the IRS, the pressure of work starts from the month March, but in the case of banks pressure on downloads starts from day 21 of each month for the payment of pensions portable to the bank; then follows another pressure from day 28 of each month, for the payment of salaries transferred to the banks; for real estate banks, as the Housing and Development Bank the pressure starts from the first of each month for payments of monthly installments; and there are many seasonal factors that may affect the change in the intensity of work, for example, the pressure to buy stocks and real state is in the summer more than during winter.

4- Changes in various sources of data: In cloud computing and storage, and input or output, where information is from public sources through servers, the capacity of computing through the CPU is limited. For example, some business downloads may need to use more than the CPU. Downloads as e-mail need a large part of the storage area and use less of the CPU , thus it is possible to buy devices with more storage capacity, but the disadvantage is that it may not be economical.

5- Unconfirmed growth models: Forecasting future needs computers, and time to supply effective servers.

How a system for managing downloads can be designed to work throughout the day in banks

There are times at which we can have a quick access to the personal accounts in banks when the number of customers, who enter to their accounts via the Internet, is small; at other times there will be high pressure on the devices.

We can determine the times at which the work downloads are certain, and therefore the pressure on the servers is significantly high, such that it does not allow for a large number of real customers to access their personal accounts. Customers are in need of confirmed or specific times to access their accounts when the pressure on servers is low. It may be needed to develop a system through studies, to determine the working time; to distribute working downloads with banks as follows:

1- From 8:30 P to 10:00 in the morning, in this period, the bank begins to open its doors to the public, and the pressure on hardware ranges from 10% to 15% of downloads for the rest of the day.

- 2- From 10:00 am to 12:00 at noon, in this period the pressure is 5%, and 20% for the rest of the day.
- 3- The time from 12:00 pm to 2:00 pm, in this period, the pressure is between 30% and 40% of downloads for the whole day.
- 4- From 2:00 pm till 5:00 pm, in this period, at the end of the work day, the pressure is from 35% to 45% of downloads for the whole day.

Work downloads are also high in certain periods during the month, for example, starting on the 20th of each month after receipt of salaries transferred to the bank and pension customers rush to their payments.

For Construction And Housing Bank, in particular, the pressure on the bank at the beginning of each month, as customers crowd to pay monthly installments for their housing or real estate they have bought.

The bank sends a bouquet of banking services provided through the Information Center, in preparation for its treatment, at this stage, the distribution of work is determined, for work download analysis during the month, which can be divided as follows:

- 1- Peak times for entering data.
- 2- Peak times for the extraction of information.
- 3- Peak times for the treatment or modification of data.
- 4- Specific data collection for work loads of in preparation for statistical analysis and reports to senior management.

Then, the bank staff distributes the tasks throughout the day so as not to block their network as a result of downloads build up at any time, which disrupts the work.

Management of banks has to determine the following:

- 1- What is the necessary time it takes to complete each task?
- 2- The distribution of all business downloads for banking services offered through the bank.
- 3- Determine the timings that are preferable to do some jobs rather than others.
- 4- Distribution of business downloads according to steps, so as the majority of the organization's operations are carried out in the times of not so great load.
- 5- The bank has to make periodic analysis of performance to determine the services that have the highest demand at specific times.

Of the primary economic advantages of cloud computing, is its ability to determine the change in the exploitation of resources that are provided through various factors affecting the cost of cloud computing; there are two important economic advantages, namely:

- 1- Use of hard labor, which can service a larger number of customers, where each customer has to pay for management applications, and show that the labor cost associated with the modernization of management, enables the organization to solve problems faced during regular work.
- 2- Limited components of servers can be used to service a large number of customers; there is a specific amount of servers for each application.

Measuring the economic savings resulting from the use of cloud computing

To measure the economic savings, from using cloud computing, we create a model to measure the distribution of costs, according to the following:

- 1- Cost of infrastructure: Cloud computing costs of the infrastructure can be changed through changes in the cost of software, such as the cost of operating systems, and software architecture, and the architectural technique which is used in creating the hardware of cloud computing, which may need to be adjusted according to the application of cloud computing.
- 2- The costs of maintenance.
- 3- Development costs.

Economic factors affecting cloud computing

There are many causes that encourage the use cloud computing as an alternative to traditional methods. These causes may change the role of information technology in the future; and many of the methods and models of provision of technological services and career structures that will need to be remodeled for compatibility with the computing structure that can be used easily through cloud computing; these reasons include:

- 1- Cloud computing is a low-cost solution.
- 2- Cloud computing solutions are characterized by flexibility and ease of responding to external changes.
- 3- Transactions and transactions developed through information technology depends, in many cases, on applications based on cloud computing.
- 4- Business users can control work remotely through cloud computing.

Portfolio of information technology investments as a measure of information technology in banks

Structure of the portfolio of information technology (Bryan Maizlish, 2005) .

Where the distribution of costs of information technology is of four layers as follows:

- 1- Costs of infrastructure: Include infrastructure assets of hardware and software
- 2- Transaction costs: Include the costs of assets applications and software and training.
- 3- Strategic costs: Strategic costs include data and information assets, for example, the costs of education.
- 4- IT costs: the costs include the costs of facilities of the operations and services assets.

It is emphasized that investments in information technology must support the objectives of the organizations and maximizes return on investment; moreover, the allocation of resources must serve the business elements of the organization.

It also must determine how to follow up the integration process between the needs of information technology in various stages of the project and the use of cloud computing applications, which can be stated as follows:

- 1- The possibility of teleworking and mobility for both business users and professionals in information technology; mobility allows savings in the costs of travel and transport from one area to another.
- 2- Distinction between the costs of infrastructure, as an asset, and acquiring part of it as a service, as in the following:
 - Hardware and operating systems.
 - Development costs.
 - Maintenance costs.
- 3- There is a difference between public computing and private computing, where they differ according to the portfolio of information technology applications for both of them.
- 4- Insurance costs and privacy.
- 5- The costs of information technology services, in order to be compatible with the management services in banks.

Information technology (IT) Portfolio

The development of the structure of information technology portfolio model in cloud computing systems

It is the process of distributing the costs of information technology to the different types of associated costs, such that we can maximize the added value resulting from investment in information technology,

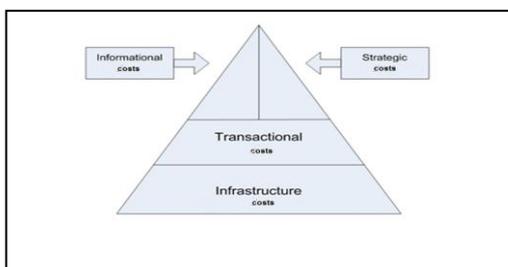
Investment in information technology must be managed as any other investment run by the organization. It is better to be managed according to the Information Technology portfolio.

IT portfolio management refers to the process of assessing and ratifying the investment of information technology as associated with other current information technology investments.

It often includes getting the right combination in most cases of investments.

Distribution of costs of information technology on the various types of cost: This can be depicted as in the following figure:

Figure (1) shows the distribution of costs in the portfolio of information technology



Rethinking in the resources of information technology

The method of distribution, of IT resources to the different classes in the portfolio, has changed. The distribution of IT costs is as follows:

Infrastructure costs. Consume from 50% to 60% of the portfolio.

Transaction costs. Consume from 20% to 30% of the portfolio

Strategic costs consume from 10% to 15% of the portfolio.

IT costs. Consume from 10% to 15% of the portfolio.

Reduction in the cost of the infrastructure became possible after recourse to the leasing of information technology resources, rather than purchased; the cost of infrastructure represented more than half of the portfolio, and thus this ratio has been reduced, and the proportion of costs, in the three other layers, is increased. Moreover, in return, the transactions costs have been reduced. The increase in IT costs that can add value to the banking service provided by a cost strategy such as innovation and change management, facilities and interaction with customers as well as increasing the proportion of cost information where possible, increased control and integration between the information, led to improving the level of banking service, and reduced the period of time spent in providing the service as well.

Low cost cloud computing solutions

Cloud computing technologies must be cost effective; if measured as part of the total cost, should improve the ratio between the costs associated with the maintenance of information technology resources, and between spending on selective IT projects.

Most of the annual budgets in many applications of IT, are spent on maintenance and repairs in return for the neglect of jobs which can provide new added value. Achieving balance is necessary, between operating costs and expenses which are incurred, in order to provide new added-value,

There is no benefit from reducing the costs of basic infrastructure, and payment of the costs of application development through the rising costs of integration between the different activities; this calls for the importance of taking another look at the real costs of information technology, which include:

- 1- Costs of integration between different functions.
- 2- Costs of Reporting.
- 3- Costs of planning for reform after crisis.
- 4- Labor costs in information technology

Costs of reducing the effects of the weaker performance of the service provider of cloud computing.

This requires more costs for insurance against the risks of cloud computing, which requires a set of skills in the field of information technology, and more support to business units. Moreover, we must compare between the costs of cloud computing and the traditional methods of information technology; while isolating various groups from each other, and reducing the resources of information technology, which must be included within business units, and seen as part of information technology.

Cost/benefit Analysis for economic evaluation

The cost-benefit analysis for the economic evaluation of information technology resources after using cloud computing is considered one of the most important methods for economic evaluation of cloud computing intangible resources. Adopting the concept of cloud computing and services by the managers of the banks, will affect the cost benefit ratio, taking into account the influence of the virtual intangible (non-physical) resources, when calculating the

benefits of investment in cloud computing. If this ratio has been compared to that calculated for the same applications that were used prior to cloud computing dissemination, we find that the proportion, of the cost of banking services, is decreased considerably in some applications, with various forms of dealing with this new technology, and with the storage space required to complete each of them. We found the following:

- 1- There is change in revenue as a result to the modification in the structure components of the hardware and operating systems of computer systems in banks.
- 2- There are specific requirements and need to provide the cash costs when creating applications based on cloud computing.
- 3- In the light of cost savings, it was possible to reduce the price of the service of e-banking, which resulted in increased demand.
- 4- It was possible to integration of work between cloud computing services, and uses of desktop computers in some applications, to ensure the privacy and security of the necessary data and to ensure improvement of the overall efficiency of banking services provided to customers

IV. THE STUDY QUESTIONS

- 1- What are the factors that cloud computing can affect the work in banks?
- 2- What are the effects expected from the change in the method of work in banks through cloud computing?
- 3- What are the future trends expected by the application of cloud computing in banks?
- 4- What are the criteria that determine the current trend towards the use of cloud computing in banks?
- 5- What are the expectations about possible outcomes from the application of cloud computing in banks?

Factors that cloud computing can affect the work in banks?

- 1- The possibility of providing new banking services to consumers: Cloud computing enabled to the provision of banking services that could not be provided through conventional technology, and is lower in cost than traditional methods, including:
 - Access to the account via the internet.
 - transferring Balance electronically
 - Payment and review of bills electronically
 - Follow-up mortgage electronically
 - Follow-up credit cards electronically
 - Follow-up lending operations electronically
 - E-banking services
 - Customer service and administration electronically
 - Sales over the internet
 - Accounts compilation
 - Electronic funds transfer
- 2- Access to more subjective services: self-computing allows the provision of self-service which facilitates operations management, and requires less participation of the workers, so as to accommodate more increase in

demand for banking services without any increase in costs.

- 3- Providing ease of movement when dealing with bank statement: it is through the follow-up of work from different places, where you can access the accounts of customers from anyplace in the world through the internet, thus facilitating the long-term negotiations, or switch in contracts easily with external contractors.
- 4- Workers gain more awareness to take advantage of the internet. : Where the workers can acquire new job skills to enable them to compete effectively.
- 5- Enables the bank to provide banking offers successfully; the bank can make banking offers as a result from savings through the use of cloud computing, as well as providing services, such as offers to deal with suppliers, as Amazon, across the internet, and buying through Google.

Reduction of technical, legal, and economic barriers that can impede the flow of banking operations, where the work is through a single network, thus contributing to the creation of new means by which to improve the working environment in the banking institution, and facilitating marketing and sales operations, and customer services. (Andropoc, 2011)

Expected effects as a result of change in working methods in the banks through cloud computing.

Cloud computing usage alters the method in which the workers use information technology resources, and how to stockpiling their data through the use of information technology resources as a virtual computer that can run complex computations independent of the ordinary one.

Expected future trends due to the application of cloud computing in banks

Cloud computing can provide an opportunity for banks to use technical information technology resources more effectively, whether to buy or lease cloud computing resources.

Criteria that determine the trend of using current cloud computing in banks

Of the most important criteria that determine the trend of using the current cloud computing in banks, is flexibility in the sense that the establishment of economical flexible system for banking services can enable fulfillment of banking services, through the multiple supply sources of services, according to client needs; this can be achieved through the costs effectiveness and productivity, that may be relatively acceptable motive to reduce the cost.

Expectations about possible outcomes from the implementation of cloud computing in banks

Cloud computing will lead to other opportunities to the banks to build models of providing services to customers with more flexibility; which leads to the growth of profitability; as a result of training the staff to deal with the technology of cloud computing.

Field study**Research objectives**

- 1) Assessment of factors that can affect the economic situation of the banks due to the application of cloud computing systems
- 2) To determine the relative economic benefit expected from the system through the application of cloud computing
- 3) Developing the structural model of the portfolio of information technology in cloud computing system.
- 4) To determine the need for modifications in the method of resource management systems in banks after the application of cloud computing.
- 5) To Learn how to design an information system that can maximize the economic benefit of the available assets, after the application of cloud computing.
- 6) Determine the potential risks of inadequate management of cloud computing systems in banks.
- 7) To identify the factors leading to the increase of economic resources in banks due to the application of cloud computing systems.

The nature and phenomena of the research problem

The phenomena of the research problem are the potential risks of inadequate management of cloud computing systems in banks. Furthermore, it is desirable to examine different methods of development of banks, so as to increase the economic resources as a result of application of cloud computing.

The research problem

Reconnaissance study showed that the application of cloud computing systems in the banks is better from the viewpoint of costs, compared to traditional methods, but it is offset by the ease of exposure of information to insurance risks. We can identify problems associated with using cloud computing application systems in banking transactions in the following points:

- 1- There are many problems relating to the use of cloud computing application systems in banking transactions due to the ease of exposure to theft of information through networks, and the subsequent identity theft.
- 2- The use of the cloud computing systems applications in banks can include exposure to many financial risks, which may lead to lower performance.
- 3- There are many natural risks resulting from the difficulty of the use of external cloud computing when dealing with customers from the outside the bank, just like buying from a commercial site as the Amazon site, but it is to take advantage only of the work during the internal computing information technology systems and unnecessary applications and the least likely to decrease in costs through cloud computing although it reasonable and cannot be ignored, but is accompanied by exposure to many risks.
- 4- **Economic savings through cloud computing in the banks faced many challenges, including:**
 - The difficulty of pricing for cloud computing services, is where there is difficulty in the development of a specific mechanism for determining the price of your resources through the participation of cloud computing and the accompanying performance, which is based on

the economic assessment of resources of information systems.

- Difficulties in securing information, particularly in the public cloud computing, in the light of cloud computing.
- Increasing the administrative costs of the information system under the cloud computing because of the large number and complexity of the procedures for data security.

Accordingly, the focus of the research problem can be in:

"The presence of several problems facing the banks as a result of difficulty in achieving a balance between the economic savings resulting from the application of cloud computing, and the ease of exposure to many risks when using external computing, in particular in banking transactions"

The importance of research

The importance of research on the application of cloud computing in banks can be summarized in the following :

1. Lack of interest by employees who work in Egyptian Banks to identify the risks associated with the application of cloud computing systems in banks.
2. The need to Identify the methods of cost savings, and to reduce the rates of exposure to risk, as a result from the application of cloud computing systems in banks.
3. The need to study how to maximize the savings associated with the implementation of a cloud computing environment versus a traditional infrastructure in banks .

Research Hypotheses

Based on the research problem and its importance, the research hypotheses could be stated as follows:

Hypotheses

- 1- There is no significance relationship between the cloud computing systems application in banks, and banking transactions, and the success of the banks.
- 2- There is no significance relationship between the application of the cloud computing systems in banks, and the fulfillment of economic savings in banks.

Research Methodology

- 1- Compiling of the theoretical framework of the search by reviewing references and scientific journals, which discussed the research topics.
- 2- Holding personal interviews between employees in the research sample of banks to know the method of work, and the implementation of programs that are planning to use cloud computing application systems.
- 3- Preparation of questionnaire form and directing it to a sample of workers in banks, where the study sample was selected on the basis of individuals who are able to give correct information in the field of study to ensure the safety of hypotheses of testing.

The research relied on the random sampling method to collect the required information from the research community because the selection of the sample is not

subject to certain conditions, due to the similarity of the working conditions in various Egyptian banks, so the researcher determined that community in Egyptian banks

The research sample

The random sampling method was applied in the field study, where the selected random sample of the study includes the following banks:

1. Bank of Egypt Iran Development
2. Bank of Alexandria
3. National Bank of Egypt
4. Housing and Development Bank
5. National Bank of Egypt
6. National Development Bank

Testing Hypotheses

To test the hypothesis of the research, the correlation coefficient test was used, to test the relationship between the variables studied, we took the average of the responses to each part of the sample of the existing organization in the research sample, and calculated this average to the nearest whole number, and then we loaded from the questionnaire form represented by each answer employee organization, then data were analyzed using statistical program SPSS / PC 17, And the degree of significance was 1%, and if the test result was significant, it meant that the null hypotheses was refused and the alternative hypothesis for both hypotheses was accepted.

First hypothesis

There is no significance relationship between the cloud computing systems application in banks, and banking transactions, and the success of the banks.

Null hypothesis

There is no significance relationship between the cloud computing systems application in banks, and banking transactions, and the success of the banks.

The alternative hypothesis

There is a significance relationship between the cloud computing systems application in banks, and banking transactions, and the success of the banks.

To test this hypothesis non parametric parameters has been used because the data are not parametric because they are the result of a questionnaire form as well as the parametric coefficients correlation was used to confirm the results, in order to study the relationship between the total answers to questions in the first group, and the total answers to questions in the second group, to see whether the relationship between them is significant or not? The results were as follows:

- 1- Parametric Pearson's correlation coefficient was used, which was equal to 0.485 at the 0.01 level of significance 0.000, meaning that the probability of error equal to zero per thousand.
- 2- Non- Parametric Correlation coefficient Kendall, was used which is equal to 0.434 at the 0.01 level of significance 0.000, meaning that the probability of error equal to zero per thousand.

- 3- Non- Parametric Correlation coefficient Spearman was used, which is equal to 0.487 at the 0.01 level of significance 0.000, meaning that the probability of error equal to zero per thousand.

Which means that there is a relationship between two variables equal to 48.5% according to the Pearson correlation coefficient, and 43.3%, according to Kendall's correlation coefficient, and 48.7% according to the Spearman correlation coefficient, and the correlation of this relationship is direct, because a correlation coefficient is positive, and the relationship is viewed as a significant relationship, that can be introduced.

Therefore we accept the alternative hypothesis and reject the Null hypothesis.

The Second hypothesis

There is no significance relationship between the application of the cloud computing systems in banks, and the fulfillment of economic savings in banks

Null hypothesis

There is no significance relationship between the application of the cloud computing systems in banks, and the fulfillment of economic savings in banks

The alternative hypothesis

There is significance relationship between the application of the cloud computing systems in banks, and the fulfillment of economic savings in banks

To test this hypothesis non parametric parameters has been used because the data are not parametric, because they are the result of a questionnaire form as well as the parametric coefficients correlation was used to confirm the results, in order to study the relationship between the total answers to questions in the first group and the total answers to questions in the third group, to see whether the relationship between them is significant or not? And the results were as follows:

- 1- Parametric Pearson's correlation coefficient was used, which is equal to 0 at 0.517 level of significance 0.01 0.000, meaning that the probability of error equals to zero per thousand.
- 2- Non- Parametric Correlation coefficient Kendall, was used which is equal to 0.430 at the 0.01 level of significance 0.000 meaning that the probability of error equal to zero per thousand.
- 3- Non- Parametric Correlation coefficient Spearman was used, which is equal to 0.487 at the 0.01 level of significance 0.000 meaning that the probability of error equal to zero per thousand.

Which means that there is a relationship between two variables equal to 51.7% according to the Pearson correlation coefficient, and 43%, according to Kendall's correlation coefficient, and 48.7% according to the Spearman correlation coefficient, and the direct correlation of this relationship, because a positive correlation coefficient, and the relationship is viewed as a moral can be introduced.

Therefore we accept the alternative hypothesis and reject the Null hypothesis.

V. FIGURES AND TABLES

As it is clear from the field study and through the open questions, that limit the opportunities provided by the use of cloud computing in the banks as follows:

Table showing the opportunities provided by the use of cloud computing in banks, and repetitions of the answer in the open question no. (1) in the questionnaire form

serial	Opportunity	Repetition
1	Providing the possibility of completing global trade transactions to accounts from anywhere in the world	10
2	Providing an opportunity for employees to work from home, which reduces the cost of labor	5
3	Increasing the speed of transactions completion by accessing the servers, and implementation of transactions.	3
4	Reducing the possibility of jobs and tasks that does not add value to the bank through the follow-up for all transactions made through the bank through internal cloud computing and excluding what is not necessary to them.	2
5	Facilitating communications and contacts with customers through the Internet, which save a lot of time spent in answering customer inquiries?	1
6	Cloud computing lead to the completion of the bank tasks more quickly as they reduce the cost associated with the use of information technology.	1
7	The Possibility to follow-up the customers through the Internet to the promotions introduced from the bank.	1
8	Enabling the application of green computing technology, by reducing the large number of computers used, also through savings in electrical energy and thus achieving green computing.	1

As is clear from the field study, and by limiting the open-ended questions that defects associated with the use of cloud computing in the banks as follows:

A table showing the disadvantages associated with the use of cloud computing in the bank and repeat the answer in open question no. (2) in the survey form

serial	disadvantages	Repetition
1	Cloud computing cannot reduce the capital costs in the existing budget.	9
2	It is difficult for employees to recognize how to avoid the risks associated with the external computerization of banks.	6
3	The emergence of non-flexible arrangements in making work in banking as a result of the presence of many of the barriers that are placed before making any bank transactions to secure the necessary data from theft, where there is more than one password its key to ensure the user accessing to their accounts before the completion of any contract to ensure the difficulty of entering hackers to any system electronically.	2
4	Many of the information technology resources are relatively high price according to the prices in 2011 for what was expected from the savings achieved through the application of cloud computing.	1
5	Although reduced costs which is provided through the use of a smaller number of computers, as well as reducing the losses in the computers, as well as reducing the losses in the exploitation of the capacity of full computers, but the insurance procedures in Banking cloud computing takes longer than the traditional roads	1
6	Many banks staff is not well qualified to deal with new technology to ensure the full benefit of them.	1
7	Increasing the probabilities of theft from banks through the ease of e-banking credit card fraud	1

As is clear from the field study and by limiting the open-ended questions that the solutions proposed to overcome the disadvantages associated with the use of cloud computing in the banks as follows:

A table showing the proposed solutions to overcome the disadvantages associated with the use of Cloud Computing in banks, and the repeated answer in open question No. (3) In questionnaire form

no	The proposed solution	Repetition
1	Reducing the transaction, that is done through external computing in banks as much as possible.	2
2	Trying to achieve economic savings through savings in operational costs such as maintenance cost for hardware or labor cost.	1
3	Rehabilitation of workers in banks to deal with cloud computing in order to be well qualified to deal with new technology to ensure the full benefit of them.	1

VI. CONCLUSION

Results:

- 1- The Integration between the savings in the cost of a cloud computing, both from the supply side or demand side that could eventually lead to significant economic savings . .
- 2- There are many risks that can occur in banks as a result of the application of external cloud computing.
- 3- The biggest concern in Investment services And banking is that the computing environment of cloud is not secure or flexible enough to enable them to bear the demands on the banking operations computerized by customers on an ongoing basis in the sense that it must confirm that clients are receiving banking services, after the signing of doing it in person during each period of time and signing it Personally with the responsible Department of the Bank.

Recommendations

Based on findings resulted from the study it can be recommended for the following:

- 1- The Use of Cloud Computing in the banks would have to be cautiously reducing the probabilities to cloud computing risks.
- 2- Making cloud services delivery model that ensures Meeting clients' Needs with Work Needs at the same time.
- 3- Taking into Account Matters pertaining to the confidential and secure data and commitment Obligations to customers and to achieve the various standards Of The quality of banking services according to the use of cloud computing.

Results and recommendations for the research

Serial	The result	Recommendation for it
1	Cloud computing technology have led to the change in the method of providing banking services in banks, so in order to comply with these developments accompanying the emergence of the need and therefore the need to apply new standards for economic evaluation of the assets after the application of that technology.	Dissemination of technology to take advantage of cloud computing and trying to lay the foundations for the economic evaluation of resources in line with the use of such technology.
2	It is No longer appropriate to use traditional methods of resources management of information technology, as doing internal transactions within the bank.	The transition to the use of cloud computing technology in all internal transactions within the bank.
3	Most of the methods used to assess the economic costs associated with the neglecting expense to secure data in transactions using cloud computing within the bank.	Update methods used in the economical assessment, including guaranteed to be inclusive of all cost items.
4	Lack of awareness from employees of many of the scientific methods needed to be done when making the work in the use of cloud computing within the bank.	Developing technical skills of workers through making training courses for them in this area.
5	The neglecting of workers to use all the possibilities available to them through cloud computing.	Informing Workers with the importance of their use of the possibilities available to them through cloud computing.
6	It may result from the use of cloud computing, many risks associated with the theft of vital data.	Further secure the vital data and save them from being stolen, lost or damaged.

ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. The author wishes to express her gratitude to all the people who support her in order to accomplish the research, and all those who contributed to the emergence of this work properly. And wishes for all success and prosperity

REFERENCES**Journal Papers:**

- [1] Paul Pocatilu, Felician Alecu, Marius Vetrici (2010). - Measuring the Efficiency of Cloud Computing for E-learning Systems - Journal of WSEAS TRANSACTIONS on COMPUTERS Issue 1, Volume 9, January 2010

Books:

- [2] Gillam, L. (2010). Cloud Computing: Principles, Systems and Applications. Springer.
- [3] John W. Rittinghouse, J. F. (2009). Cloud Computing: Implementation, Management, and Security. CRC Press.
- [4] Harms, R., & Yamartina, M. (novemeber 2010). The economics of the cloud. Microsoft press.
- [5] Bryan Maizlish, R. H. (2005). IT Portfolio Management Step-By-Step: Unlocking The Business Value of Technology. John Wiley & sons.
- [6] Miller, M. (2008). Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online. Que Publishing
- [7] Rajkumar Buyya, J. B. (2011). Cloud Computing: Principles and Paradigms. John Wiley & Sons.
- [8] Toby Velte, A. T. (2009). Cloud Computing, A Practical Approach. McGraw-Hill Prof Med/Tech.,
- [9] Tim Mather, S. K. (2009). Cloud security and privacy: an enterprise perspective on risks and compliance . O'reilly Media, Inc.

Proceedings Papers:

- [10] Androcec, D. (2011). Research Challenges for Cloud Computing Economics. CECIIS. Central European Conference on Information and Intelligent Systems, CECIIS - 2011.
- [11] aatun, M. G. (n.d.). Cloud Computing: First International Conference, CloudCom 2009, Beijing, China, December 1-4, 2009, Proceedings. Springer, 2010.

E-Payment: Prospects and Challenges in Nigerian Public Sector

Peter M Ogedebe¹ Phd, Babatunde Peter Jacob² MSc.

^{1,2}Computer Science Department, Faculty of Science and Technology, Bingham University, Karu, Nasarawa State, Nigeria

Abstract: This study presents E-payment system as means of fast tracking the implementation of government policies through the elimination of delays in government payment system and minimizing interactions between government officials and contractors to eliminate opportunity for corruptive tendencies. This is with the view of achieving economic and efficient transactions in government finances and improve quality of reporting system in the Nigerian public sector. Though the e-payment system is faced with challenges, like public acceptability, lack of uniform platform being operated by the banks, lack of adequate infrastructure and issues of security, with the proper use of the E-payment system, corruption which is a cancer in the public sector will be adequately addressed.

Keywords: Corruption E-Payment, Nigeria, public sector,

I. Introduction:

Trade normally involves exchange of goods and services and equivalent abstract value such as money [1]. Money has been in use as an abstract way of representing value and a system for making payments. With development, new and abstract way of representing value were introduced [2]. They also observed that “a corresponding progression of value transfer system starting from barter through bank notes, payment order, cheques and later of credit cards has finally culminated in electronic payment system”. A payment system is a system used for transferring money. It is a system that utilize cash substitutes, traditional able payment systems are negotiable instruments such as draft (eg cheques) and documentary credits such as letter of credit.

The advent of computers and information technology has made it possible to have alternative electronic payment systems like debit cards, credit cards, electronic fund transfer, direct credits, internet banking and e-commerce payment systems. Payment systems may be physical or electronic and each has her own procedures and protocols. Generally speaking, e-payment can refer to a payment system for buying and selling goods or services offered through the internet or any type of electronic fundtransfer.

Nigeria was drifting towards anarchy because of the high level of corruption at all levels in public sector of the nation. Worried by this trend, the President during the presentation of the 2009 Budget to the National Assembly announced a number measures aimed at addressing this cankerworm. One of the measures was the unveiling of e-payment regime, this involve directing that, all financial transactions of the Federal Government be made electronically effective from 1st January 2009.

Many people see e-payment in different perspectives. To the contractor, he is only interested in payment for the services he has rendered immediately the services are

delivered and certified. The civil servant is only interested in the prompt payment of his or her salary at the end of the month.

But academically and generally speaking, what do we really understand by the term “e-payment”? Some expert described e-payment as a subset of e-governance which is the application of electronic means of effecting payments by government to civil servants and contractors. It is a form of direct payments and banking without physical appearance at the Ministries, Departments and Agencies (MDA) or bank through the means of electronic interactive communication channels and other technology infrastructure [3].

E-payment according to Agba [4] is a method of effecting payments from one end to another through the medium of the computer without manual intervention beyond inputting the payment data.

II. Statement of the problem:

There are unacceptable delays in the payment of government contractors in Nigeria. Most of these delays are as a result of bureaucracy within government circle. The high level of corruption in the public sector due to the interaction between the government contractors and government officials is another major issue of concern. Most government officials usually ask for gratification to facilitate payments. Payment of workers’ salaries is usually delayed because of the same issue of bureaucracy. E – Payment offer many advantages over the usual manual payment. Some of the advantages include, making payment swiftly and remotely, easy tracking of payment to beneficiaries account to assist in audit trail. These often proof very helpful to corruption fighting agencies like Economic Financial Crime Commission (EFCC) and Independent Corrupt Practices Commission (ICPC) in cases of investigation.

III. Objectives of the study:

The study specifically identifies the following objectives [5] and highlight how e-payment can help to eliminate the following:

1. The use of cash in government transactions in order to speed up payment for all government activities.
2. Minimize the interactions of government officials and contractors in order to reduce or eliminate corrupt tendencies.
3. Many risks that is associated with carrying large sum of money, these include threat from armed robbers, fraud, thefts and others.
4. Enhance real time reporting and improve quality of financial reporting system in the public sector.
5. Achievement of economy and efficiency in government financial transactions.

IV. Types of e-Payment

Dankwambo [5] stated that there are two types of e-Payment in the Nigerian context. These are:

i. End to End Processing

Here, all the processes from approvals to the receipt of value by the beneficiary are done electronically.

ii. Manual e-Payment or use of Mandate.

It is the mixture of manual and electronic process where the available infrastructures cannot support the End to End processing.

There are however many forms of e-payments, these include cards, internet mobile payment, financial services kiosks, biometric payments, electronic payment networks [2]. Many of these payment systems have become globally available; these include credit card and Automated Teller Machine (ATM) are other specific form of payment for financial transactions for products in the equity market, bond markets, currency markets, future markets, derivative markets, option market and for transfer of funds between financial institutions both domestically using clearing and Real Time Gross Settlement (RTGS) system and international issuing SWIFT network [6]

An ideal E-payment means that the Ministries, Departments and Agencies (MDA) issuing payment instruction(s) follow an electronic process within establishment. The payment moves electronically from desk to desk for approval before it gets to the bank and approval(s) also must be given electronically. Instructions and then sent to their banks electronically who in turn effect to all the banks of which the accounts of their beneficiaries are domiciled from the comfort of their office. Thereafter, associated schedules are immediately made available to third parties receiving the payment. Thus, they are able to view all account balances across banks on one screen, monitor the status of all instructions and are able to see why any instruction has not been effectively carried out.

However, what is presently in operation is "manual e-Payment". The government's road map shows clearly that the ultimate method is the "End -to- End Processing". When will Nigeria attain the goal or road map? Only time will tell!

V. Prospects

An ideal e-Payment system should possess the good point as identified by Dankwambo [3].

- Easy tracking of payments to beneficiaries' account hence it will assist audit trail.
- It reduces cases of corruption
- It will assist corruption fighting agencies like the EFCC and ICPC in cases of investigation.
- It is the beginning of a cashless society
- Overall increase in the efficiency of operation
- Reduced transactions of very low value
- Increase convenience of payments
- Payment can be made swiftly and remotely using various devices
- Accountants will appreciate IT more and this will improve the quality of financial reports generated by MDAs.
- Economic Growth and Development as Transparency and Accountability improve.
- Real Time Reporting and

- Eliminate writing of cheques

The risk associated with cheques been stolen, forging of signature and disparity between amount in words and figures has been totally eliminated.

VI. Challenges

The problems militating against e-Payment as listed by Sumanjeet [7] generally revolve around.

- Integrity: to ascertain that transmitted financial information is unchanged in transit.
- Non-repudiation: to ascertain that all parties have non-deniable proof of receipt.
- Confidentiality: to ascertain that transactions are protected from possible eavesdroppers.
- Reliability: to ascertain that there is reduced possibility of failure.
- Authorization: to ascertain that individuals are recognized and granted the desired rights and privileges.

Public Education and Acceptability

The system which is still in its stage requires a lot of information and education of the public to enable them appreciate the laudable programme put together by government to protect their interests. The banks also need to be carried along in the implementation. If they are properly and adequately educated. Then the chances of the total acceptable of the programme can be assured. Furthermore, many see e-Payment as an imposition.

6.1 Lack of Uniform Platform by Banks and MDAs

There is no compelling law mandating the banks to use common software platform. Every bank is left to use whatever platform that they felt will perform the e-Payment services on behalf of the clients. There is the problem of switches in effecting transfer from one bank to another. Interconnectivity has been a problem. No uniformity of account numbers since different banks different numbering systems. Happily enough, the Federal Government according to Dankwambo through the Office of Accountant General of the Federation will be rolling out a common platform configuring soon.

6.2 Lack of adequate Infrastructure

As noted earlier, the e-Payment system is being partially implemented. If it is to be fully implemented, a number of IT infrastructures will have to be put in place. These include but not limited to Laptop, desktop, scanners, good internet connectivity, training and global software. The provision of basic Information Technology infrastructures according to Ovia [8] is a major challenge.

6.3 Platform Security

As rightly pointed out by Atanbasi [9], the major challenges of e-Payment in the country is security. Security in terms of platform, hackers and virus attacks. This will ensure that output from the system are reliable and accurate. The MDAs still carry their schedule(s) to the banks through the banks through Compact Disks (CDs), Flash Drives or e-mail attachments.

6.4 Lack of Seriousness by Banks

While a number of banks have deployed the necessary infrastructure in place to ensure effective implementation, it is sad to note that some banks are still not fully ready for this new payment regime.

6.5 Resistance to changes in technology among customers and staff due to:

- Lack of awareness on the benefits of new technologies
- Fear of risk
- Lack of trained personnel in key organizations
- Tendency to be content with the existing structures, and
- People are resistant to new payment mechanisms
- Security- where disclosed of private information, counterfeiting and illegal alteration of payment data may be rampant [3].

Omogui-Okauru [10], noted some of the problems with e-Payment at the Federal Inland Revenue Services (FIRS), among others, have to do with reconciliation; being able to reconcile what is paid, what the banks received and ultimate basic account. The complaints and constraint facing accountant in FIRS is also the fact that the e-Payment has not been as fast as should be.

VII. Evaluation of E-Payment in Nigeria

Looking at the objectives for introducing e-Payment in the public sector in Nigeria, one can state very categorically that only few of the objectives have been met. These are:

- Eliminate many risks associated with carrying large sums of money such as armed robbery, fraud theft and others.
- At least government organizations no longer pay cash to “contractors” and civil servants.
- Elimination of the use of cash to facilitate speedy payments for all transactions. But to a very large extent, the following objectives have not been met.
- Fast tracking the implementation of government policies through the elimination of delays in government payment system. There are still instances of delay in payment to contractors who are not ready to play ball. There has been a complaint from some contractors handling projects in the rural areas over difficulties associated with the e-payment model.
- Minimize interaction of government officials and contractors to eliminate opportunity for corruptive tendencies. It will be difficult to eliminate this as interaction at which ever level will continue formally or informally if Nigerians are to be honest with themselves. In which ever case, there is need to ask the question. Who are the contractors? Is the due process of government working or not? Who are the officials subverting this and other laudable programmes of government? Can corruption really be stamped out of the system?
- Achievement of economy and efficiency in government financial transactions. For as long as corruption remains within the polity, there can be no efficiency in the system. The EFCC and the judiciary will have to find a common ground to tackle this cankerworm that has defiled all solution. China’s example could be the

best solution. However, this may also not work because of religious and tribal sentiment among some Nigerians.

- Enhance real time reporting and improve quality of financial reporting system in the public sector. It has been observed that since the implementation of the policy, there have been late returns or no response in respect of unapplied funds. The existing system cannot guarantee real-time reporting of finances. As result there can be no good financial reporting.

VIII. Conclusion,

E-payment exist in various forms. These include the use of cards, internet mobile payment, financial services like kiosk, end to end processing and the use mandate. Its proper implementation will help to speed up payments in government transactions, fast track implementation of government policies by eliminating delays in payment, reduce corruption by minimizing interaction of government officials with contractors. The system also has the prospect of helping to achieve economic and efficient government financial transactions and enhancement of real-time reporting and improvement of quality of financial reporting system in the public sector. Whatever may be the shortcomings of this system thus far, it is better started than waiting to evolve an efficient system before putting the system into use. The government deserves kudos in its efforts to rid the society of corruption. The expectation of Nigerians on the war against corruption is no doubt very high.

References

- [1] Sadeghi, A., Schneider, M., "Electronic Payment Systems", in Becker, E., et. al., (eds.), *Digital Rights Management*, (2003), p.121.
- [2] T O Asaolu, T J Ayoola & E Y Akinkoye. Electronic Payment System in Nigeria: Implementation, constraints and solutions. *Journal of Management and Society*.1 (2), 2011, pp. 16-21
- [3] I. HDankwambo, Understanding the e-payment system. Paper at the workshop on In Abuja 30th March 2009
- [4] D. Agba, (2010), Implications and challenges of E-payment System. Retrieved from <http://www.itnewafrica.com>. [Accessed on 20th November, 2010]
- [5] O. I Osibote, , E-payment System: Processes, Procedures, Challenges and Prospects. Paper at the workshop on In Abuja. (2010). Accessed on 5th May 2010 from. <http://financial.tmcnet.com/mergers-acquisitions/news/2010/10/18/5073144.htm>
- [6] Payment system. Retrieved from http://en.wikipedia.org/wiki/Electronic_payment.
- [7] S. Sumanjeet. Emergence of payment systems in the age of electronic commerce: The state of art, *Global Journal of International Business Research* 2(2), 17-36.
- [8] J Ovia, Payment System and Financial Innovations. A paper presented at the Annual Policy Conference, Nov. 2002
- [9] L A Atabansi. We can grow our Technology. Retrieved from: <http://theeconomyng.com/interview16.html> (Accesed on 10th July, 2012).
- [10] I Omogui-Okauru, FIRS Calls for improved E-payment system by Director General of the Federal Inland Revenue Board. Retrieved from <http://www.vanguardngr.com/2010/03/firs-calls-for-improved-e-payment-system> (Accessed 10th July 2012)

Modified-DES and Triple Modified-DES Encryption Algorithm with Reduced BER in Wireless Communication

¹B. Divya, ²S. M. K. Sukumar Reddy,

¹M. Tech Student, Department of Electronics (E.C.E), Vaagdevi Institute of Tech & Science, Peddasettipalli(Village), Proddatur(M.D), Kadapa(Dist), Andhra Pradesh A.P, India.

ABSTRACT: In this paper, we propose a modification to the existing Data Encryption Standard (DES) to make it secure and prone to the bit errors caused by the wireless channel. We observe that using the modified algorithm in wireless channels, improves the bit error rate (BER) performance as well as security compared to DES. Wireless channels are an open medium to intruders and their attacks; encryption is a vital process to assure security over these channels. However, using well-known encryption algorithms to encrypt data in wireless communication will result in a catastrophic error due to the avalanche effect, which is implemented in these algorithms to assure security. Although this effect is desirable to assure security, these algorithms do not take into account the bit error characteristics of the wireless channel. So the need for a new secure encryption algorithm that takes into account the bit error characteristics of wireless channels becomes necessary.

Keywords: Avalanche effect, Encryption, security, cryptography, bit error rate

I. INTRODUCTION

Encryption is an essential process to assure confidentiality over wireless channels, because wireless channels are an open medium to intruders in which they can intercept and alter the contents of any transmitted information. Well known standardized encryption algorithms such as DES and AES were designed to achieve security against intruders. Hence, DES and AES were designed in such a way to satisfy the avalanche effect criteria. The avalanche effect of an algorithm requires that changing a single bit of the key or the plaintext (i.e: the data before encryption) to change half the bits of the ciphertext (i.e: the data after encryption) on average [1]. This effect also requires that a single bit change of the ciphertext or the key will result in some significant and random looking changes at the plaintext. On average half the bits will be in error. Any algorithm that has this property does not exhibit any statistical correlation between input and output that an adversary might use in attack, thus the algorithm will multiply bit errors, that is if there is one error at the received ciphertext, there will be many errors at the decrypted plaintext. However, this effect becomes catastrophic in wireless channels, because wireless channels tend to add noise to the signal. So the wrong reception of a single bit in a certain block at the receiver will result in half the bits of the decrypted block on average to be in error. So it is clearly noticed that an algorithm that satisfies the avalanche effect is very sensitive to bit errors. If one bit of the ciphertext is received in error, then each bit of the plaintext will have 0.5 probability of error. This means that there is a trade off between the security and the bit error rate. So encrypting the information with traditional encryption approaches will significantly degrade the performance of wireless networks,

especially when the SNR is too bad. Some researchers had noticed this problem and tried to solve it in two main approaches. In [1] the authors introduced what is called opportunistic encryption, in which they encrypt the data with longer keys which implies more security whenever the SNR of the channel is higher so that the probability of error at the received ciphertext will be lower than for lower SNR values and hence higher security can be used. They also used forward error correction (FEC) codes to protect encrypted packets from bit errors. They also assumed perfect knowledge about the channel in order to use opportunistic encryption. Using their new encryption technique, the authors showed that the throughput of the system was more utilized than for that of using fixed encryption of AES. In [2] the author introduced a new mode of operation in which the data is transmitted, the design goal of the author work was to design a mode of operation for encryption algorithms that has substantially less error propagation than other modes. The Author called the new mode ECFB; the new mode has a lower bit error rate than other modes. However, the BER was not significantly improved. Another way of thinking about a solution for this problem was to design new algorithms which are specific for wireless applications. Some encryption algorithms such as MISTY1, KASUMI and KASUMI-R were specifically designed for wireless applications such as the Universal Mobile Telecommunications System (UMTS) [3]. However, these algorithms were only designed with security in mind. In [4], the authors show that these algorithms satisfy the avalanche effect as in other traditional encryption algorithms.

These algorithms are also shown to be vulnerable to different types of attacks as shown in [5]. Now it becomes clear that a new algorithm which takes into consideration the error nature of wireless channels is critically needed. In this paper, we propose a new algorithm to be used in wireless communication.

The proposed algorithm is a modification to DES encryption algorithm by which we are achieving a very high security compared to that of DES. The proposed algorithm is modified to have much lower bit error rates than that of DES. We studied the effect of each S-Box in DES on the BER. We redesigned the S-Boxes so that the error is reduced. We also introduced a new round to the sixteen rounds of DES, to increase the security of the algorithm to an unbreakable level and to reduce the bit error rate significantly. In this paper, the new algorithm is still a 64-bit input algorithm as DES, but with 128-bit ciphertext and 136-bit key. The rest of the paper is organized as follows: Section II describes the architecture of the proposed algorithm (M-DES) and Triple M-DES. In Section III the performance of the

proposed algorithm is evaluated. Simulation results are presented in Section IV. Finally, some conclusions are drawn in Section V.

II. Proposed Modified Algorithm Architecture

In this paper, we introduce two main modifications to the standard DES in order to improve the BER performance of the received data, and to enhance its security. Fig. 1 shows the general architecture for the proposed modified- DES (M-DES). As shown in the figure, the first sixteen rounds in the proposed algorithm have the same structure as the standard DES rounds [6], except that in the MDES we propose using different S-Box mapping tables than that of standard DES. All of the the first four S-boxes in the M-DES will have the same mapping table as the mapping table in the first S-Box of original standard DES, while the rest of the S-Boxes (number five through eight) will have the same mapping table as the mapping table of the second S-Box of standard DES. The S-Box is a mapping table that maps a 4-bit input to 6-bit output. In standard DES, each of the eight S-Boxes has a distinct mapping table. The S-Boxes were initially designed in such a way to meet the avalanche effect criteria [7].

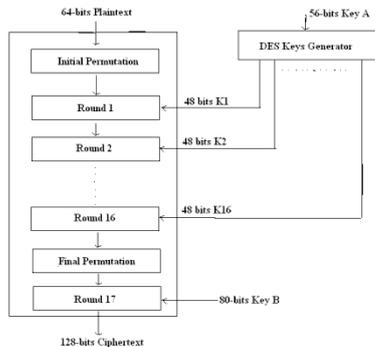


Fig. 1: M-DES General Architecture with the addition of Round 17, which has two inputs and one output. Inputs are 64-bit output of final permutation and 80-bit key, the output is a 128-bit ciphertext

In the MDES, we are still having eight S-Boxes in each round, but with only two distinct mapping tables rather than the eight distinct mapping tables in standard DES. This in fact will improve the BER performance, however, at the price of reducing security. In our proposed algorithm, we will also propose another modification not only to compensate for this reduction in security but also to improve security as compared to standard DES. This will be described in the rest of this section.

Our motive for the second modification is in fact coming from the work in [8], where the authors showed that DES can be cracked using the differential cryptanalysis attack if the attacker has 2^{47} pairs of plaintext and ciphertext. The authors also show that each distinct S-Box mapping table requires around $2^{6.5}$ pairs to be cracked. Therefore, in our proposed M-DES, while using only two distinct mapping tables (rather than the eight distinct mapping tables in the standard DES), the number of pairs needed to crack the algorithm reduces to 2^{13} pairs. To overcome with this security reduction, we introduce round 17 in the M-DES as shown in Fig. 1. By introducing round 17, the algorithm becomes in fact secure to both brute force and differential cryptanalysis attacks as will show later in this section. Round 17 has two

inputs and one output, the two inputs are the 64-bit output of the final permutation and an 80-bit key, the output is the 128-bit cipher. The 80-bit key is used to map the 64-bit input of Round 17 to a 128-bit output. This mapping procedure is shown in Fig. 2 where the 64-bit input of round 17 is divided into sixteen sub-frames of four bits each. While the output 128-bit consists of 32 sub-frames of four bits each. Each five bits of the 80-bit key is used to map one of the 4-bit input sub-frames to one 4-bit output sub-frame. So the input sub-frames will be scrambled in 16 out of the 32 output sub-frames. The remaining 16 sub-frames of the output are randomly filled with zeros and ones.

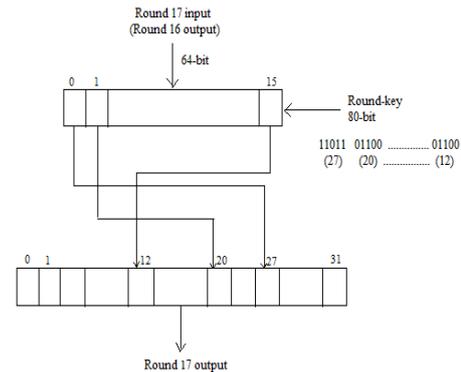


Fig 2: Round 17 Design, mapping the 16 input sub-

frames into the 32 output sub-frames using the 80-bit key. For example, if the first 5 bits of the key are 00101. This means that the first 4-bit sub-frame of the input will be mapped to the fifth 4-bit sub-frame of the output. At the receiver side, the receiver will have the 80-bit key, so he will be able to recover the useful 64-bit ciphertext out of the total 128-bit received ciphertext. So the proposed algorithm has a plaintext of 64 bits, a ciphertext of 128 bits and a key of 136 bits. So round 17 will scramble the useful 64-bit encrypted frame inside a 128-bit frame. In order for an intruder to crack the algorithm, 2^{13} useful pairs of plaintext and useful ciphertext are needed. Assuming that an intruder has the ability to encrypt 2^{13} plaintexts and get their corresponding 2^{13} ciphertexts. The intruder can't directly use those pairs to crack the algorithm, because he will need to guess the useful 64-bit out of each 128-bit cipher he obtains. So the intruder will need to try all the possible 64 bits combinations out of the 128 bits for each of the 2^{13} ciphers. The probability of getting one correct pair (64-bit input and its corresponding 64-bit output out of the 128-bit cipher) is given by

$$P_1 = \frac{1}{\binom{32}{16}} = 1.6637 \times 10^{-9}$$

(1)

So the probability of getting 2^{13} correct pairs is

$$P_2 = (1.6637 \times 10^{-9})^{2^{13}} \approx 0$$

(2)

This probability is very small, so the algorithm is considered immune to differential cryptanalysis. Round 17 also solves a very critical security problem in DES, by the new 80-bit key of round 17, M-DES will have a total of 136 bits as its key. Using the Brute force attack (testing all possible 2^{56} keys), DES was shown to be cracked in

less than 20 hours [6]. However, having a key size of 136 bits in M-DES, this problem was resolved, because it is impossible for an intruder to try all possible 2^{136} in a feasible amount of time. Assuming that the key of DES can be cracked in only one second, it will need more than 136 trillion years to crack M-DES. Therefore, M-DES is secure against brute force attack.

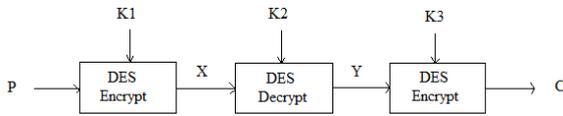


Fig. 3: Triple DES architecture

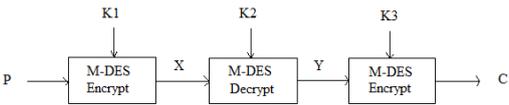


Fig. 4: Triple M-DES architecture

Similarly we can implement this concept for Triple DES (Fig.3) also. Fig.4 shows architecture for Triple M-DES. We can observe an improvement in the BER performance along with much improved security in triple M-DES.

III. PERFORMANCE EVALUATION

A. System Model

In this paper, we simulated the performance of the proposed algorithm in ModelSim using verilog. A large sequence of bits was generated randomly at the sender. The generated bits were then divided into 64-bit blocks; each block was then encrypted to a 128-bit block using M-DES. The encrypted blocks were then assumed to be transmitted over the wireless channel. The encrypted blocks received in error are then decrypted block by block. Then, the resulted sequence of bits at the receiver after decryption is compared with the sequence of bits at the sender before encryption, and the error is calculated. The same procedure is done using DES, Triple DES and Triple M-DES.

B. BER Analysis

Using simulation we show that the performance of M-DES in terms of BER is much better than DES. Initially, we studied the BER of DES, assuming one bit error only. We studied the effect of the error analytically up to the end of round three. Then it became harder to complete the sixteen rounds analytically, so we used simulations to evaluate the BER of DES, and we found that on average there was 32 bits in error out of the 64 bits. Simulations also show that in M-DES, having one or more errors at the received ciphertext block will not result in half the decrypted bits to be in error, while for DES if any bit is received in error, this will result in half the bits to be in error.

C. Security

We showed in details in the previous section how the algorithm is considered secure against brute force and differential cryptanalysis attacks. The addition of the new 80-bit key, make it impossible to crack the algorithm using brute force attack. While the addition of the new round reduces the probability of cracking the algorithm using differential cryptanalysis to almost zero.

D. Key Management

An efficient way to share the new key is to add the new 80-bit key to the old 56-bit key and deal with them as one key at the transmitter. Both the sender and the receiver will need to divide it into two keys when encryption or decryption is done.

E. Complexity

The proposed algorithm is less complex than DES in terms of the number of distinct S-Box mapping tables. However, the number of rounds and the key size of M-DES is more than DES, which might increase the complexity compared to DES.

IV. Simulation Results

This section provides the simulation results of the BER obtained for M-DES and Triple M-DES compared to DES. Here encrypted text is the output of encryption algorithm and encipher text is the input to the decryption algorithm after transmission. Here we have considered a random error bit is the changed bit while transmission. Here i and j are the iteration variables used for the 64 bits and 100 iterations respectively. The average ber is calculated which is observed as 31.62 for an ordinary DES algorithm, as shown below (Fig.5):

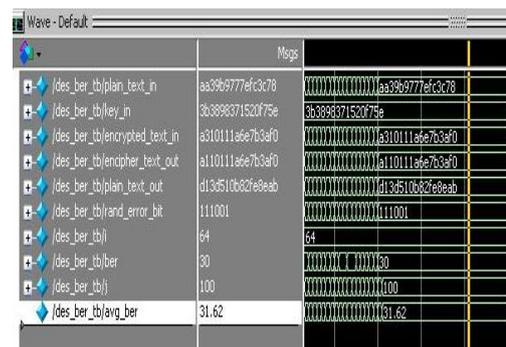


Fig.5: Average BER output for standard DES algorithm

Similarly, for M-DES algorithm requires the same 64-bit plaintext as for standard DES, but it takes a 136-bit key and produces a 128-bit ciphertext. Here, the second key is the 80-bit key which is given to the Round 17. Here, we are assuming that the encrypted text differ from the encipher text by one random bit. So, here the average ber output is 18.16, which is comparatively much less than standard DES algorithm. Thus the ber performance is improved. (Fig.6)

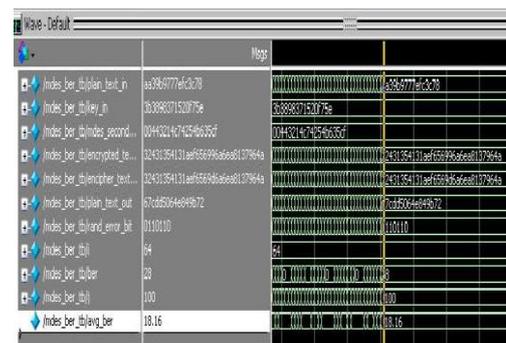


Fig.6: Average BER output for M-DES algorithm

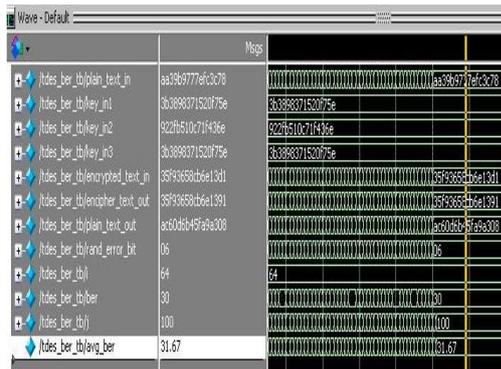


Fig.7: Average BER output for Triple DES algorithm

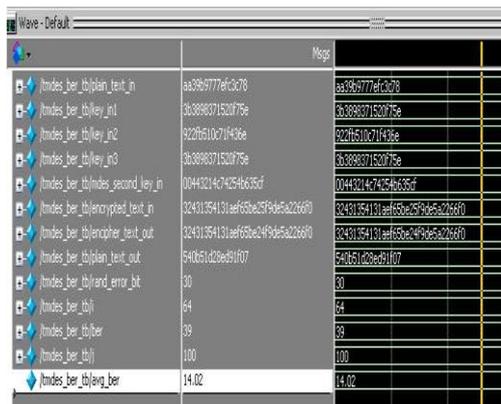


Fig.8: Average BER output for Triple M-DES algorithm

Similarly, we can observe the results for the triple DES algorithm (Fig.7); in this the DES algorithm is used three times. So, the average ber calculated for triple DES is 31.67, which is almost same as for the standard DES.

Similarly, if we use M-DES in triple DES, i.e if we use three M-DES instead of standard DES than the average ber is obtained as 14.02(Fig.8). This is better than normal triple DES. Thus the ber performance is much more improved and the security is also increased using triple modified DES.

V. CONCLUSION

In this paper, we showed how well known block cipher encryption algorithms are not sufficient for the use in wireless applications. Because in these applications, the signal may experience severe degradations and attenuations, which causes wrong reception of the signal and hence a catastrophic effect on the decryption process. In this paper, we proposed a new encryption mechanism based on modified DES. Using simulation we quantified the performance of the proposed M-DES versus the known standard DES and Triple DES versus Triple M-DES for encryption in a wireless communication channel. We showed that the new algorithm outperforms the standard DES and Triple DES algorithms in terms of error performance. We also showed that the new algorithm enhanced the security to a high security level which is prone to all applicable types of attacks.

REFERENCES

- [1] M. Haleem, C. Chetan, R. Chandramouli and K.P.Subbalakshmi, "Opportunistic Encryption: A Trade-Off between Security and Throughput in Wireless Networks," IEEE Transaction on Dependable and Secure Computing, vol. 4, no. 4, pp.313-324, Oct 2007.
- [2] Reason, "End-to-End Confidentiality for Continuous-Media Applications in Wireless Systems," PhD dissertation, UC Berkeley, Dec.2000.
- [3] "Technical specification group services and system aspects," 3GPP TS 55.216 V6.2.0, September 2003.
- [4] Sedat Akleyek, "On the avalanche effect of MISTY1, KASUMI and KASUMI-R," Master's thesis, Middle East Technical University, Feb 2008.
- [5] O. Dunkelmann, N. Keller and A. Shamir, "A Practical-Time Attack on the A5/3 Cryptosystem Used in Third Generation GSM Telephony," Cryptology ePrint Archive, Report 2010/013, Feb 2010.
- [6] Behrouz A. Forouzan, Cryptography and Network Security, 1st Edition, Mc Graw Hill, 2008.
- [7] M. Matsui, "On correlation between the order of S-boxes and the strength of DES," in Proceedings of Eurocrypt94 (A. De Santis, ed.),no. 950 in Lecture Notes in Computer Science, pp. 366375, Springer-Verlag, 1995.
- [8] E.Biham and A.Shamir, "Differential Cryptanalysis of the Full 16-round DES," Proceedings of Crypto'92, vol. 740, Santa Barbara, CA, December 1991.

Electric destroyer of anopheles' larvae for vector control in malaria

Hangnilo Robert

Laboratoire d'Électrotechnique, de Télécommunication et d'informatique Appliqué, LETIA ; Ecole Polytechnique d'Abomey-Calavi, EPAC, Université d'Abomey-Calavi, UAC ; 01BP2009 Cotonou-Bénin.

Abstract: The emergences of insecticide resistance in *Anopheles* and drug resistance in plasmodium have become a matter of great concern for malaria control in Africa South of Sahara. So many solutions have been applied except any electromagnetic challenge. Electromagnetic field can be fatal for all living being. In larval stage *Anopheles* are very vulnerable. We learn from *Malapterurus electricus* catfish the mastering of a tool that imitates the electric organ discharges (EOD) which the catfish uses to kill victims in fresh water. The given tool devastates *Anopheles*' larvae massively and spontaneously whatever their number. Hot water of more than 42 °C only can kill mosquitoes' larvae, but the electric field destroys them even at 26 °C for the temperature of the water in the tool.

Keywords: Malaria, electric organ discharge, *Malapterurus electricus*, vector control, electric destroyer

I. Introduction

Malaria disease burden is a real disaster for Africa. There is nowhere in Africa where one can stay and be shielded from malaria havoc, Fig. 1- map 1. In Tunisia there has been no local transmission of malaria since 1979. Although the disease has been eradicated, its re-emergence is not to be considered impossible [1].

Malaria disease burden continues to be a major impediment to health and wealth in Sub-Saharan Africa. Every year that disease kills in Africa a lot of young children and pregnant women, [2], [3], [4], [5]. Because malaria is such a common disease and well known to the people it affects most, and because most of those who become sick with malaria do not visit health care facilities, assessing the size of the problem, and how it is changing over time, is an enormous challenge, [6], [7].

Roll Back Malaria program was launched as a catalyst for a renewed worldwide commitment to face up to a disease that has been ignored so long a time by the world. That strange sickness is putting a heavy brake on development, particularly in Africa. Malaria remains the single biggest cause of death of young children in Africa and one of the most important threats to the health of pregnant women and their newborns, [8].

World Health Organization, (WHO) defines malaria control as reducing the malaria disease burden to a level at which it is no longer a public health problem. The objectives of malaria vector control are two-fold: to protect people against infective malaria mosquito bites by reducing the vector longevity, the vector density and human vector contact; and to reduce the density of local

malaria transmission at community level, and hence the incidence and prevalence of infection and disease.

As a consequence of all there has been tremendous progress in increasing access to Insecticide Treated Nets, (ITNs) particularly in the Republic of Benin.

Antimalarial drug resistance becomes a major public health problem which hinders the control of malaria. Study of those who are treated for uncomplicated malaria demonstrated clinical failure for chloroquine and for Pyrimethamine-Sulfadoxine, (PS). Two issues raise serious doubts about the use of PS as a replacement for chloroquine. The former is detection of persistent asymptomatic parasitemia at day 7 in children treated with PS. The latter is widespread use of one PS component (sulfadoxine) to prevent opportunistic infection in Acquired Immunodeficiency Syndrome, (AIDS) patients or to treat infection, [9], [10], [11],[12].

The other side of the drama is the insecticide resistance in malaria. The emergence of resistance in *Anopheles* to common class of insecticides was reported from many countries in Africa. That resistance affects the major vectors of malaria: *An. gambiae* and *Anopheles funestus* s.l. In several parts of Africa the resistance of pyrethroid insecticides in *An. gambiae* is going to be investigated, [13]. The vector resistance to pyrethroid insecticides affects the malaria situation in some endemic areas of Africa by reducing the efficacy of ITNs as well as of Indoor Residual Spray, (IRS). If nets are not timely replaced, they eventually deliver a sub-lethal pyrethroid insecticide dose that selects for resistance, [14].



Figure 1. Malaria and electric catfish in Africa

Source: 1-: Recommendations canadiennes

2- Mao

Antimalarial policy despite its voluntarism and significant progress has not brought the last blow from which the health of people in humid tropical countries can be considered permanently spared of serious nuisance due to malaria disease burden. Thus, in spite of the great deal of effort deployed to control malaria disease burden, the target seems to remain as a whole.

However we have not to give up. It is quite possible to win malaria. We know that electric field can be fatal for every living being. The evidence is given by *Malapterurus electricus* catfish which naturally hunts in fresh water by the means of an electric pulsed wave. *Malapterurus electricus* counts in the strongly electric intermittent catfish in the world. It is a typical catfish of Africa, [15], Fig. 1-map 2.

The target of our paper is to present the results of a study which concerns the destruction of *Anopheles*' larvae in the way of *Malapterurus electricus* catfish. We have set up in this purpose the model and the equivalent electric scheme of that catfish, [16]. This tool delivers electric waves in the way of the catfish and destroys spontaneously the larvae of *Anopheles* in the water. We can now hope the reduction of *Anopheles*' populations in the endemic areas for malaria.

II. Materials And Methods

The electric organ discharge of electric catfish is produced by some tissues equipped with electric charges that can create an electric field and potential, [17]. In practice we need a capacitor to create an electric field. To prove the feasibility of a destroying of anopheles' larvae by electric discharge, we deal first with a parallel-plates capacitor, Fig. 2-a. Its capacitance is given by Eq. 1:

$$C_1 = \frac{A\epsilon_0}{d} \tag{1}$$

Where *A* is the surface of each plate, *d* is the distance between the plates and $\epsilon_0 \approx 8.85 \cdot 10^{-12}$ Farad/m stands for the permittivity constant for the vacuum. The efficacy of the so conceived tool is impressive. But the tool possesses some weakness concerning its short range and the eventual corrosion of the iron made frames of the capacitor. As the tool must stay in water and cover a long distance, we found the solution given by two copper wires, Fig. 2-b. The capacitance of two cylindrical wires is given by Eq. 2:

$$C_2 = \frac{2\pi\epsilon_0\ell}{\ln\left(\frac{d}{a}-1\right)} \tag{2}$$

Here, *a* stands for the radius of each cylinder, *d* for the distance between the axis of the cylinders and *l* for the length of the wires. With a copper wire the weapon will not corrode in fresh water and can be spread out any required distance. To test the weapon in Laboratory we made a wooden artificial river, Fig. 3. *Anopheles* larvae can live in sweet water as well as in brackish one. Thus we have tested the weapon in each quality of water respectively and have got similar results.

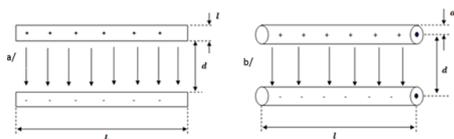


Figure 2. a- Capacitance of two parallel-plates capacitor
 b- Capacitance of two cylindrical wires



Figure 3. The anopheles' larvae destroyer in a wooden lake

[18] and [19]. A simulation in Matlab of the potential difference of two points in space gives a twofold straightening alternation curve. That remark has enabled us to set the electric equivalent scheme of *Malapterurus*

Electricus up as a Graetz bridge rectifier: that is the simulator of electric discharge of the catfish. We made a parallel-plates capacitor with iron plates of area $S = 4677, 54 \cdot 10^{-6} \text{ m}^2$ separated a distance $d = 10, 05 \cdot 10^{-3} \text{ m}$, Fig. 2-a. We put the capacitor in a non conducting waterproof material and drop in some water with *Anopheles*' larvae, and then we connect the capacitor to the power supply through the simulator of electric discharge. After a moment we notice that all the larvae are destroyed. The electric weapon we are setting up would be a real challenge if we could spread it out a long distance in fresh or brackish water. We have also succeeded in destroying larvae with the capacitance of two copper wires of 1, 5 mm section to reach the goal, Fig. 2-b and Fig. 3. For every test we measure the temperature of the water. We took *Anopheles*' larvae from *Sainte Cecile* in Cotonou a typical endemic area for malaria in the Republic of Benin.

III. Results and Discussion:

The parallel-plates capacitor is iron made and cannot perform for a long time in water as a result of corrosion. We found the capacitance of two copper wires as the providential solution, (see Fig.2-b). The copper will not corrode and the wire can cover any required distance. As we can see on Tab 1 and 2 larvae destruction power depends on the current nature and the value of the voltage applied to the simulator: with a pulsed voltage of 90.70 V a magnitude we need the same duration (01 min 22 sec 49 ties) to destroy larvae as it is in the case of direct voltage but the power delivered in pulsed voltage mode is lower than the direct voltage one. In each case the power rises with the voltage value and the destructive duration decreases in the same time, (Tab 1 and 2, Fig. 4.). The heating temperature of the water also increases with the voltage, see Fig. 5. To save energy we must set up a weakly pulsed voltage electric weapon to imitate nature perfectly. In fact we can realize now why every electric catfish gives out solely pulsed EOD, see Fig. 1-

map 2: nature works wonders and always saves energy. Hot water can kill anopheles larvae provided only that its temperature goes beyond 43° C. We have heated separately water and we measure its temperature and we drop in larvae. Our observations are recorded in Tab 3

Table 1
 Power delivered by the pulsed wave generator

Magnitude, V	Time of destruction, mn:ss:tt	Courant, mA	Puissance, W
90.70	01:22 : 49	1.9	0.172
107.67	00:50:01	2.5	0.270
116.15	00:43:10	2.9	0.337
130.30	00:31:01	3.1	0.404
144.44	00:21:96	3.6	0.520
155.75	00:19:23	3.9	0.607
167.06	00:13:75	4.2	0.702
178.38	00:12:12	4.5	0.803
186.86	00:11:22	4.7	0.878
198.18	00:09:22	5.2	1.030
206.66	00:08:11	5.3	1.096
212.32	00:06:80	5.8	1.231

ve characteri

Table 2 Electric current field

Power delivered by the direct voltage, V	Time of destruction, min: sec: tie	Courant, mA	Power, W
90.70	01:22:49	3.1	0.281
107.67	00:50:01	5.9	0.635
116.15	00:43:10	6.5	0.755
130.30	00:31:01	6.8	0.886
144.44	00:21:96	7.4	1.069
155.75	00:19:23	8.8	1.371
167.06	00:13:75	10	1.671
178.38	00:12:12	10.6	1.891
186.86	00:11:22	11.5	2.149
198.18	00:09:22	13.1	2.596
206.66	00:08:11	14.1	2.914
212.32	00:06:80	13.8	2.930

Table 3 Larvae killing with hot water

Température of water °C	Larvae are:	Alive = 0
		Numbered = 1
		Dead = 2
45.0	2	
43.2	1	
41.5	0	
40.0	0	
39.2	0	
38.5	0	
36.6	0	

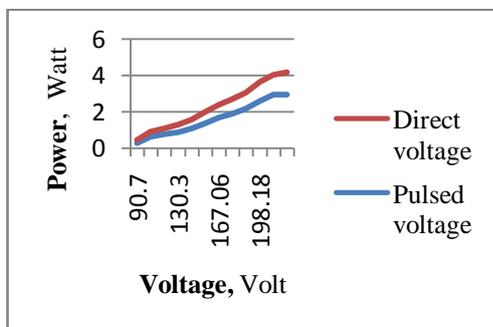
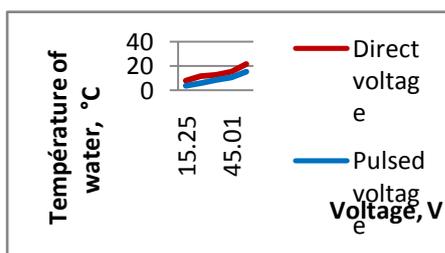


Figure 4. Power delivered by each type of electric field

The weapon destroys larvae even at 28, 6° C (with 90, 70 V and 15 sec 25 tie in duration) whereas hot water of more than 43° C only kills larvae: hence the electric weapon is the only one responsible for the *Anopheles*' larvae destruction.



5. Figure Temperature behaviour of the water

IV. Conclusion

Our paper is about to achieve its aim. The objectives of malaria vector control are well known: protection of people against infective malaria mosquito bites by reducing the vector longevity, the vector density and the human vector contact; and the reduction of the density of local malaria transmission at community level...

Our work focuses directly on a major point of malaria vector control' objectives that concerns vector density reduction. It would be more effective to control *Anopheles*' larvae than mosquito that got wings to fly. Larvae are more concentrated in fresh or brackish water than *Anopheles* is in the air. Mosquitoes' larvae are generally vulnerable in electric field but *Anopheles*' larvae are the most. The reason is as follows: when culex's larvae (for example) come up the water' surface to catch oxygen, they keep themselves diagonally in the water whereas *Anopheles*' larvae lie in a plate at the water surface to catch oxygen and simultaneously a potential difference if there is any. The true vector of malaria disease burden is *Anopheles*' larvae. On top of indoor strategy of control by ITNs and IRS we have to fight against *Anopheles*' larvae which are the primary vector of malaria. The tool of such a new strategy of vector control is the electrostatic weapon that we have set up in our Laboratory. The weapon devastates spontaneously all larvae located inside its sensibility in fresh or brackish water.

References

- [1] Chahed, M K; Bouratbine, A; Krida G, Ben Hamida, A. (2001): Réceptivité de la Tunisie au paludisme après son éradication : analyse de la situation pour une adéquation de la surveillance. Bull Soc Pathol Exot, **94** (3), pp. 271-276.
- [2] Mouchet, J *et al.* (1997): The reconquest of Madagascar highlands by malaria. Bull Soc Pathol Exot, **90**(3), pp. 162-8.
- [3] Mouchet, J *et al.* (1991) : Le défi de la lutte contre le paludisme en Afrique tropicale : place et limite de la lutte anti vectorielle. Cahier d'études et de recherches francophones/ Santé, **1**(4), pp. 227-88.
- [4] Migliani, R; *et al.* (2003): Le paludisme vu des tranchées : le cas de la cote d'ivoire en 2002- 2003. Med Trop., **63**, pp. 282-286.
- [5] Louis, J P *et al.* (1992) : Le paludisme-maladie dans la ville de Yaoundé (Cameroun) : prise en charge et lutte anti vectorielle au niveau familial. Bulletin de la Société de pathologie exotique (1), pp. 26-30.
- [6] Julvez, J; Mouchet, J; Ragavoodoo, C. (1990) : Epidémiologie historique du paludisme dans l'archipel des Mascareignes (Océan Indien). Ann. Soc. Belge MM Trop, **70**, pp. 249-261.
- [7] Bayouhd, F *et al.* (1995): Le Paludisme Dans Les Troupes Onusiennes en Somalie lors des Operations Humanitaires "Restore Hope" et "Unosom". Médecine du Maghreb, **54**, pp. 19-22.
- [8] Mouchet, J; Coosemans, M. (1991): Quelles structures pour une lutte anti vectorielle ? Ann. Soc. Belge Med Trop., **71** (1), 249-261.
- [9] Henry, MC; Niangue, J; Kone, M. (2002) : Quel médicament pour traiter le paludisme simple quand la

- chloroquine devient inefficace dans l'Ouest de la Côte d'Ivoire. *Med. Trop*, pp. 62:55-57.
- [10] Baudon, D; Martet, G. (1997): Malaria and travelers: Protection and information. *Médecine tropicale*, 57 (4 bis), pp. 497-500.
- [11] Danis, M. (2003) : Avancées thérapeutiques contre le paludisme en 2003. *Med Trop*, pp. 63 : 267-270.
- [12] Desfontaine, M *et al.* (1990): Evaluation des pratiques et des coûts de lutte anti vectorielle à l'échelon familial en Afrique centrale. II Ville de Douala (Cameroun). *Ann. Soc. Belge Med. Trop* 70, pp. 137-144.
- [13] Yadouleton, A W M *et al.* (2009): Development of vegetable farming: a cause of the emergence of insecticide resistance in populations of *Anopheles gambiae* in urban areas of Benin. *Malaria Journal*, 8:103.
- [14] N'Guessan, R *et al.* (2007): Reduced efficacy of insecticide-treated nets and indoor residual spraying for malaria control in Pyrethroid resistance area, Benin. *Emerging infectious diseases*, 13(2), pp. 199-206.
- [15] Rankin, CH and Moller P. (1992): Temporal patterning of electric organ discharges in the African electric catfish, *Malapterurus electricus*. *J Fish Biol* 40 (1), pp. 49-58.
- [16] Hangnilo, R and Adanhounme, V. (2012): Modeling the electric organ discharge of *Malapterurus electricus* catfish. *IJSER*, vol. 3, Issue 3, August-2012.
- [17] Schwann HP. (1994): Electrical properties of tissues and cell suspensions: mechanisms and models. *Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Baltimore, MD, USA.
- [18] Hangnilo, R and Adanhounme, V. (2012): Solving the oscillating electric dipole equation by the Adomian decomposition technique. *IJSER*, vol. 3, Issue 5, May-2012.
- [19] SA., Adédjouma, R. Hangnilo, JM. Zonou, GA. Mensah, "Determination of the electric energy characteristics produced by the *Malapterurus electricus*, an electric catfish of fresh water in Benin", *Revue Cames-Série A*, 12 (1), pp. 52-57, 2011.