International Journal of Modern Engineering Research (IJMER)

Editorial Board

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

Kunjal 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

Design of Operational Trans conductance Amplifier in 0.18μm Technology

SHWETA KARNIK [#1], AJAY KUMAR KUSHWAHA [#2], PRAMOD KUMAR JAIN [#3], D. S. AJNAR[4],

[#1, 2, 3,4] Micro Electronics and VLSI design

*** Electronics & Instrumentation Engineering Department, SGSITS, Indore,(M.P.)India**

Abstract- This paper presents design concept of Operational Transconductance Amplifier (OTA). The 0.18μm CMOS process is used for Design and Simulation of this OTA. This OTA having a bias voltage 1.8 with supply voltage 1.8 V. The design and Simulation of this OTA is done using CADENCE Spectere environment with UMC 0.18μm technology file. The Simulation results of this OTA shows that the open loop gain of about 71 dB which having GBW of 37 KHz. This OTA is having CMRR of 90 dB and PSRR of 85 dB. This OTA having power dissipation of 10 mW and Slew Rate 2.344 V/μsec.

Keywords- OTA, Cadence, CMRR, PSRR, Power Dissipation, CMOS IC Design.

1. INTRODUCTION

Due to recent development in VLSI technology the size of transistors decreases and power supply also decreases. The OTA is a basic building block in most of analogue circuit with linear input-output characteristics. The OTA is widely used in analogue circuit such as neural networks, Instrumentation amplifier, ADC and Filter circuit. The operational Transconductance Amplifier (OTA) is basically similar to conventional Operational Amplifiers in which both having Differential inputs. The basic difference between OTA and conventional operational Amplifier is that in OTA the output is in form of current but in conventional Op-Amps output is in form of Voltage.

This paper is organized as follows. Section II describes brief description about operational Transconductance Amplifier (OTA) design. Section IV describes Simulation Results of OTA. Section IV describes the conclusion of this paper.

2. OPERATIONAL TRANSCONDUCTANCE AMPLIFIER (OTA) DESIGN

Figure 1 shows the schematic diagram of Operational Transconductance Amplifier (OTA). In this OTA the supply voltage is VDD= 1.8V. In the below circuit of OTA the Transistors, M11_1 and M11 work as a constant current source and Transistors M1, M2 and M3 works as two current mirror 'pairs'. The Transistors M4, M5, M6 and M7 are the Differential amplifier.

The Transistor M8 is an output amplifier stage. The design parameters of this OTA are shown in below table I.

There are several different OTA's are used in which this OTA is a simple OTA with low supply voltage and high gain. The Op-amp is characterized by various parameters like open loop gain, Bandwidth, Slew Rate, Noise and etc. The performance Measures are fixed Due to Design parameters such as Transistors size, Bias Voltage and etc. In this paper we describe design of OTA amplifier and this design is done in 0.18μm technology.
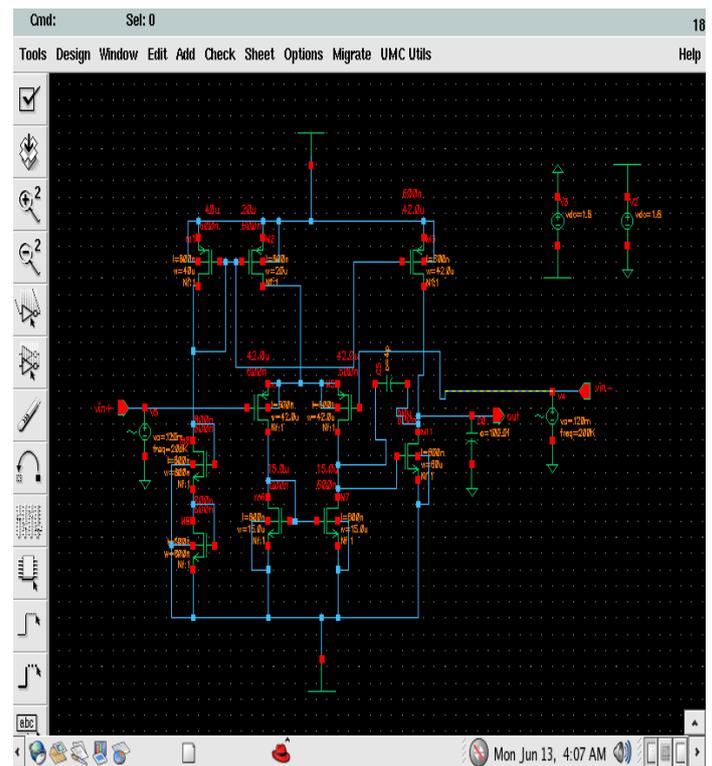
Figure 1: Operational Tran conductance Amplifier

TABLE I
TRANSISTOR SIZE

Device	W/L(µm)
M1,M2,M3	40/0.6
M4,M5	20/0.6
M6,M7,M8,M9	42/0.6
M8,M9	50/0.6
M10,M11	60/0.6
M12,M13	0.8/0.6
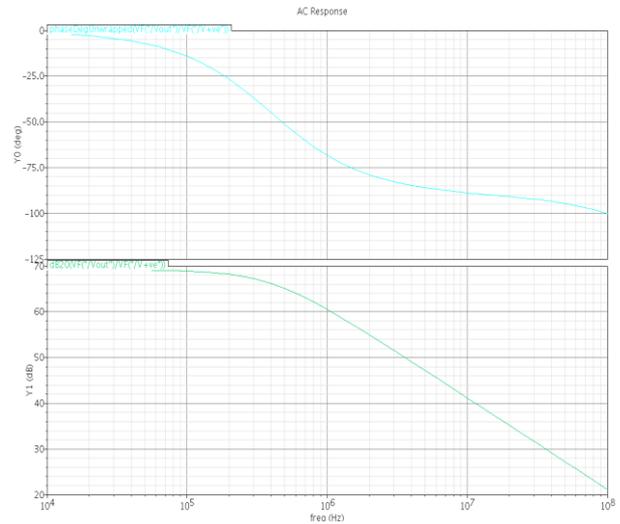
Figure 2: Shows AC response which shows gain and phase change with frequency.

3. SIMULATION RESULTS

The design of this Operational Transconductance Amplifier (OTA) is done using Cadence Tool. The Simulation results are done using Cadence Spectre environment using UMC 0.18 µm CMOS technology. The simulation result of the OTA shows that the open loop gain of approximately 71 dB. The OTA has GBW of about 37 KHz.

The Table II shows that the simulated results of the OTA. The AC response which shows gain and phase change with frequency is shown in figure 2. Figure 3 shows the DC sweep response of This OTA. The Transient response with input in pulse is shown in figure 4. Figure 5 illustrates PSRR variations with frequency. The variation in CMRR is shown in figure 6.

The simulated results of this OTA shows that PSRR of 85 dB and CMRR of 90 dB.
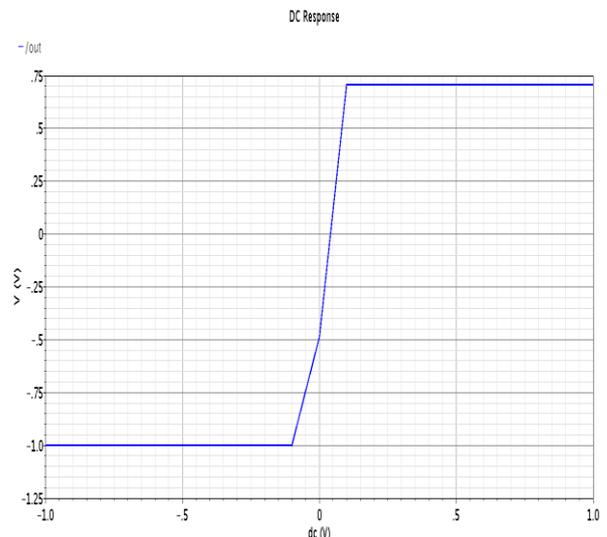
Figure 3: DC sweep response.

TABLE II
SIMULATED CHARACTERSTICS OF OTA

Specifications	Simulated
CMOS technology	0.18µm
Open loop gain	71 dB
Supply voltage	1.8 V
Bias Voltage	1.8V
PSRR	85 dB
CMRR	90 dB
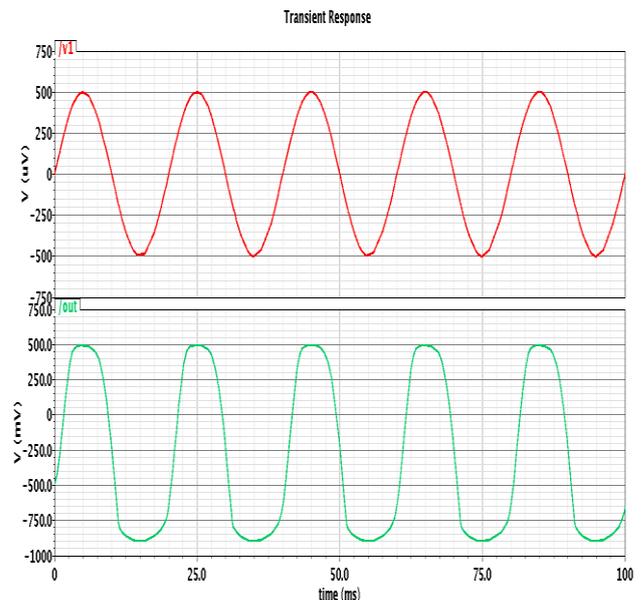
Figure 4: Transient response with input is pulse.
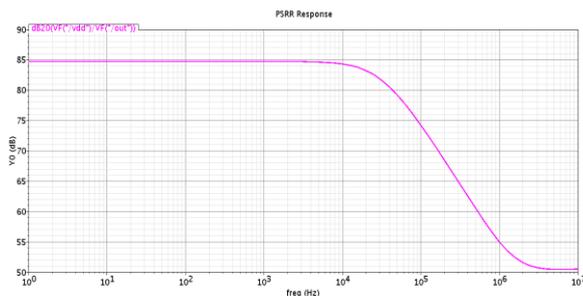
Figure 5: PSRR change with frequency.
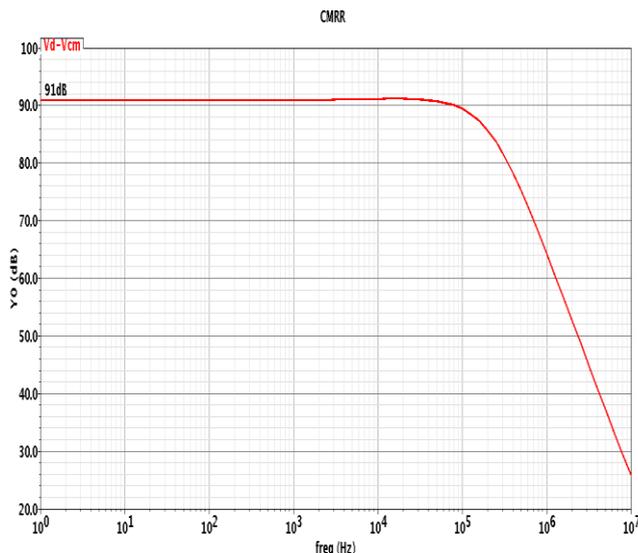
Figure 6: Change in CMRR with frequency

4. CONCLUSION

In this paper we present a simple Operational Transconductance Amplifier (OTA) topology for low voltage and low power applications. This OTA can be used in low power, low voltage and high time constant applications such process controller, physical transducers and small battery operated devices. This work can be used in filter design, ADC design and instrumentation amplifiers because of its high gain, high CMRR and low power consumption.

5. REFERENCES

[1] J. H. Botma, R.F. Wassenaar, R. J. Wiegerink, "A low voltage CMOS Op Amp with a rail-to-rail constant-gm input stage and a class AB rail-to-rail output stage", IEEE 1993 ISCAS, Chicago, pp.1314-1317.

[2] Paul R. Gray, Paul L.Hurst, Stephan H.Lewis and Robort G.Mayer "Analysis and design of analog integrated circuits",Forth Edition, John Wiley & sons, inc.2001, pp.425-439.

[3] Adel S. Sedra, Kenneth C.Smith "Microelectronic Circuits", Oxford university press, Fourth edition ,2002,pp.89-91.

[4] Jin Tao Li, Sio Hang Pun, Peng Un Mak and Mang I Vai "Analysis of Op-Amp Power-Supply Current Sensing Current-Mode Instrumentation Amplifier for Biosignal Acquisition System",IEEE conference,August-2008,pp.2295-2298.

[5] Y. Tsividis, Operation and Modeling of the MOS Transistor, 2nd ed. Boston, MA: McGraw-Hill, 1998.

[6] D. A. Johns and K. Martin, Analog Integrated Circuit Design. New York: Wiley, 1997.

[7] Phillip E. Allen and Douglas R. Holberg "CMOS analog circuit design" ,second edition, Oxford university press, 2007,pp. 269-274.

Authors Profile:
SHWETA KARNIK

MTECH degree in Microelectronics and VLSI Design from SGSITS Indore 2012, working in the field of VLSI Design. B.E degree in Bio-Medical Engineering from Rajiv Gandhi technical university Bhopal ,INDIA in 2009.

AJAY KUMAR KUSHWAHA

MTECH degree in Microelectronics and VLSI Design from SGSITS Indore 2011, working in the field of analog design B.E. degree electronics and communication engineering from Rajiv Gandhi technical university Bhopal 2009. He is now working as Asst. Prof. in Department of Electronics & Communication Engineering, MITM, Indore (M.P.),

PRAMOD KUMAR JAIN

He has received the B.E. degree in Electronics and communication Engineering from D.A.V.V. University, India in 1987 and M.E. Degree in Digital Techniques & Instrumentation Engineering from Rajiv Gandhi Technical University Bhopal, India in1993. He has been teaching and in research Profession since 1988.He is now working as Reader in Department of Electronics & Instrumentation Engineering, S.G.S.I.T.S Indore, His interest of research in Analog and digital system design.

D. S. AJNAR

He has received the B.E. degree in   Electronics and Communication Engineering from D.A.V.V. University, India in 1993 and M.E. Degree in Digital Techniques & Instrumentation Engineering from Rajiv Gandhi Technical University Bhopal, India in 2000. He has been teaching and in research Profession since 1995. He is now working as Reader in Department of Electronics & Instrumentation Engineering, S.G.S.I.T.S Indore, India. His interest of research is in Designing of analog filter and Current conveyer.

SEGMENTATION AND COUNTING OF PEOPLE THROUGH COLLABORATIVE AUGMENTED

[1]Akhil Khare, [2]Kanchan Warke, [3]Dr. Akhilesh Upadhayay

Abstract In this System counting of people is done with the help of computer vision. The reason behind this is that computer vision is the field which is concerned with the automated processing of images from the real world to extract and interpret information on a real time basis. The image data can take many forms. For example views from multiple cameras, multidimensional data from medical scanner etc. Coding for this is developed in openCV (Open Source Computer Vision Library). This library consists of programming functions for real time computer vision and it is developed by Intel. From this library programming functions such as image preprocessing, morphology processing, image marking etc. are used for counting of people.

1. INTRODUCTION

With the rapid development of economic society, the crowd flowing in various public places and facility is more and more frequent. Effectively managing and controlling the crowd in public places become an important issue. People counting system based on this kind of demand arises, which can be used in the crowd surveillance and management, but also can be used in commercial domain such as market survey, traffic safety as well as the architectural design domain and so on. The research on counting people has the profound significance and the broad prospect because it directly or indirectly improves the staffs' working efficiency and the utilization of building facilities in various places. In the past history of this project different methods have been developed to count the number of people. But some of them have problems associated with them; hence we are trying to overcome them in this system. In developing the method for counting the number of people in complex indoor spaces, our goal is to develop a method such that it should be robust, easily realizable and effective. It should have high recognition rate in relatively stable environment and relatively sufficient light.

A people counter is a device used to measure the number and direction of people traversing a certain passage or entrance per unit time. The resolution of the measurement is entirely dependent on the sophistication of the technology employed. The device is often used at the entrance of a building so that the total number of visitors can be recorded. Many different technologies are used in people counter devices, such as infrared beams, computer vision, thermal imaging and pressure-sensitive mats.

2. LITERATURE SURVEY

Authors like Lin SF, Chao HX, T.Zhao, R. Nevatia addressed issue of people counting. It consist of methods like: fitting method based on low level feature, feature point tracking, object detection method. Fitting method is easy to use, but as it has neglected individual concept and skipped single object tracking process, it becomes difficult to acquire correct people counting information. Object tracking method

has high precision because it detects directly object. And feature point tracking method acquires people counting information by tracking moving feature point, then applying cluster analysis for further point track. But though this method is insusceptible of camera angle, but has lower accuracy.

Hence to overcome these difficulties and problems new method should be invented. And this method can be easily realized and suitable for environment. It can also be useful for understanding personal information. Author Xi Zhao, Emmanual, Dellandrea and liming Chen mensioned in paper "People counting System Based on Face Detection and Tracking in video." that in literature most of work is relied on moving object detection and tracking, based on the assumption that all the moving objects are people.They proposed an approach in which people counting is based on face detection , tracking and trajectories classification. Scale invariant Kalman filter combined with kernel based object tracking algorithm is used to handle face occlusion. They proposed a strategy to count people by automatically classify face trajectories. Then two Earth Movers Distance based classifier is used to discriminate true and false trajectories.

Due Fehr, Ravishankar Sivalingam, Osama Lotfallah, Youngchoon Park described in paper "Counting People in Groups" the importance of camera surveillance in the era of growing security concerns, and it is also necessary. They mentioned that there is successful development of detecting abandoned objects and people tracking. People tracking is relatively easy as compared to people counting in groups. Mutual occlusion is the most problematic in group counting. Several techniques for group counting estimation is suggested such as foreground \detection using mixture of Gaussian, foreground detection using pixel layering, shadow Emoval.

Duan-Yu Chen, Chih-Wen Su, Yi-Chong Zeng, Hong –Yuan Mark Liao proposed a system "An Online People Counting System for Electronic Advertising Machine" for counting the number of people watching a TV-wall advertisement or electronic

billboard without counting repetitions by using stationary camera. In this first of all face detection and face filtering is done, in which, SVM based face detector is used. Face filtering is used to filter false positive face. Then feature extraction is performed on torso of human subject. Then an online classifier trained by Fisher's Linear discriminant strategy is developed.

Fang Zhu and Xinwei Yang suggested "People Counting Based on Support Vector Machine**"** infrared people counting method. In data processing procedure pattern recognition idea can be introduced according to characteristics of time continuous data collected by infrared sensors. In this method people counting is based on two steps. First is data acquiring in which infrared signal information is collected. Second is data processing in which noise removing and normalization is done by standardization and data segmentation. Then feature extraction is performed. Lastly classification and identification of people who go through infrared area is done. When several people go through infrared signal at the same time, this method counts number of people accurately.

3. PROPOSED SYSTEM

In this paper, a new robust method for counting people in complex indoor spaces is presented. As shown in Fig.1 the method for counting people diagram, the method has counted the number of people in the indoor spaces through four modules: image pre-processing module, morphology processing module, image marking module and people counting module, in order to master the information of the indoor for increasing efficiency and utilization of building facilities. Image pre-processing module chooses image greying, background subtraction based on threshold, median filtering algorithm and threshold segmentation to eliminate background interference. The morphology processing module uses the improved erosion operation and the improved dilation operation to extract target feature. Then the following image marking module uses connected component detection algorithm, setting the object feature and shape judgment condition to remove false contouring and marking object region by rectangle frame. Finally, people counting module is used to count the number of people.

A. Image Processing Module

The captured video images need pre-processing in the method for counting people. In our method, the main function of image pre-processing module is to eliminate background interference and extract the foreground object information, that is, the foreground object in the image sequence will be extracted from the background. The result of this module as the basis of the people counting will directly affect the accuracy of people counting result. First, in image pre-processing module we capture images using a single camera, which is hanged in the middle of the roof in order to

cover the entire housing and own a better sensitivity. Secondly, we use image greying turn current image and background image into two gray images. Thirdly, we use background subtraction based on threshold process the two gray images to extract the foreground object for detecting the relative static and moving human object. Finally, we use median filtering method eliminate noise and then use maximum between-cluster variance threshold segmentation method turn the foreground object image into a binary image. Now we detail the image pre-processing module.
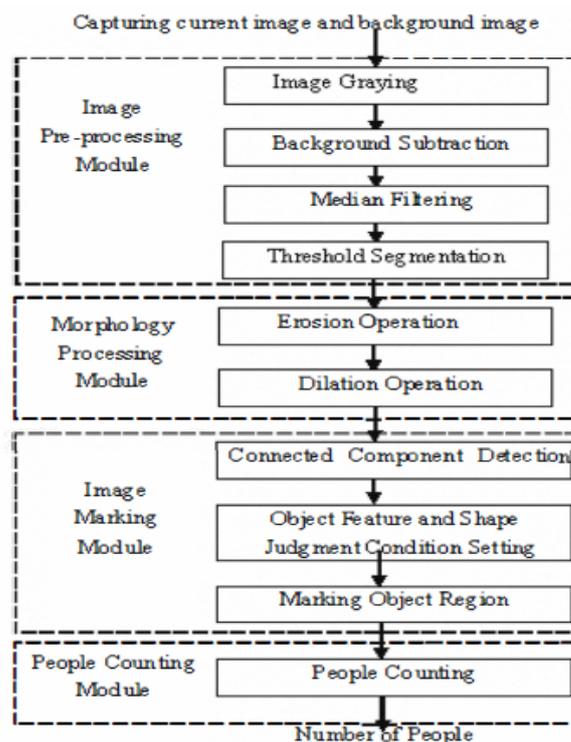
Figure 1.   The method for counting people diagram

1) Image greying

Image greying is defined by throwing away the colour information and using gray express image luminance. In the beginning of image pre-processing module, we use image greying turn the current colour image and the background colour image into two gray images. Image greying is to make the colour components R, G, B equal. Gray image has 256 Gray Levels because R, G, B range is from 0 to 255. In this paper we perform image greying thought weighted average method, which gives R, G, B different weights and makes the value of R, G, B weighted average as follow:

$R=G=B=rR+gG+bB$ (1)

Among analysis, we can gain the most reasonable gray image when r=0.299 , g=0.587, b=0.114 as follow:

$R=B=G=0.299*R+O.587*G+0.114*B$ (2)

2) *Background subtraction based on threshold*

Using image greying, two gray images which include the current gray image and the background gray image are received. We use background subtraction based on threshold process the two gray images to eliminate background interference and extract the foreground object information image.

Threshold selection is a key issue. As the gray values of head generally below 90, we choose maximum between cluster variance adaptive threshold method whose threshold is chosen within the range [0, 90]. If the pixel gray difference is bigger than the threshold, the pixel value in input gray image is seen as foreground stored in the image, else the pixel is considered as white pixel which value is 255. Through those processing, the majority of background disturbance is eliminated. Moreover, in some public spaces such as cyber bar, computer room, laboratory, the computer frame to the object extracting influence should be considered. Because computer frame and the top of head have approximate gray value, the head which locates near computer will be divided into two sections only using background subtraction. Allowing for this question, if the frame gray value of current image below 90 and the number of pixels which variation of the frame upper and lower or the frame left and right are bigger than the threshold is bigger than the set number, the pixel value in input gray image is seen as foreground stored in the result image, else the pixel is considered as white pixel which value is 255.This improved method effectively resolves the computer frame disturbance question.

3) *Median filtering method*

After background subtraction based on threshold, the foreground object images have a certain extent noise interference. The noise makes image quality deteriorated, causes the image blurred, even submerges the image feature and affects the analytic result. Therefore in the pre-processing module we adopt median filtering method to eliminate noise. Median filtering commonly uses a sliding template including the odd number of points, with the median of each template window gray value instead of the gray value of designated point. In this system, + template median filtering is used to eliminate the noise of foreground object image. After arranging the values of five pixels including the pending pixel and 4-neighbors of the pending pixel from small to big, we choose the median of the gray levels as the value of the pending pixel. Median filtering can obvious reduce noise and make image smoothing, which filters the small object blocks and highlights the feature information we need.

4) *Threshold segmentation*

Threshold segmentation is fundamental approach to segmentation that enjoys a significant degree of popularity. It needs a right threshold to divide the image into object and background. Maximum between-cluster variance threshold segmentation algorithm is used to change the object image after median filtering into a binary image. This algorithm as follows:
 In our method, threshold value T is chosen within the range [0, 90] because gray values of head generally below 90. The result of threshold segmentation is a binary image including object information.

After image pre-processing module, we receive a clear binary object image, which is eliminated background interference and beneficial to the next processing.

B. Morphology Processing Module

Mathematical morphology processing (6) is widely applied to image processing, which mainly includes dilation, erosion, opening and closing operation. Because the binary object images after image pre-processing module often have the discrete noise and holes in object region, morphology processing module is used to remove the isolated noise and fill the hole in the object region, which first uses an improved erosion operation and then uses an improved dilation operation.

1) *Improved erosion operation*

In this system, an improved erosion operation is proposed, which does the first erosion operation using 3 x3 template as B to process the binary object image, then does the second erosion operation using r template as B. Through two times erosion operation, the binary image is removed isolated noise and becomes clean.

2) *Improved dilation operation*

After erosion operation, an improved dilation operation is used in the method, which performs the first dilation operation using template as B, then performs the second dilation operation using 3x3 template as B. Through two times dilation operation, those holes in the object region are filled and some gaps are bridged.

Morphology processing module can improve the accuracy of counting system through enhancing the object feature. This step has laid a good foundation for the further image marking.

C. Image Marking Module

Image marking module aims to mark the head region. First, image marking module uses connected component detection algorithm, then sets the object feature and shape judgment condition, fmally, removes false contouring based on the object feature and shape judgment condition, simultaneously uses rectangle frame mark object region.

1) *Connected component detection algorithm*

Connected component detection algorithm[7] is to fmd all the pixels which belong to the same connected component and to give the same marking to the same connected component pixels. Through this algorithm, we gain a marking image in which the value of each pixel is the value of its regional marking. As shown in Fig.2 the image marking scheme, the connected

component detection algorithm has been done as follows:

Setting the initialization of marking counter is 0 and using column-based scan method to mark those pixels (the gray values are equal to 0)based on those marking of those pixels' four neighbor pixels which have been scanned, at the same time carrying on the following marking algorithm judgment:
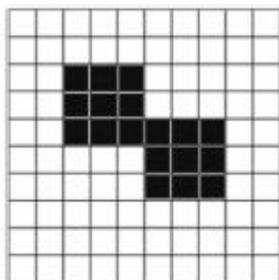

Figure 2.   Image marking scheme

Step I: If the gray values of four pixels which separately lie the lower left, the left, the upper left, the up of current pixel are 255, the marking counter adds one.

Step 2: If the gray values of four pixels which separately lie the lower left, the left, the upper left, the up of current pixel have the same marking but not all are equal to marking value 0 , the marking is given current pixel.

Step 3: If the gray values of four pixels which separately lie the lower left, the left, the upper left, the up of current pixel have different marking and two kind of marking (not including the marking is zero), judge the size of two kind of marking, the small marking is given current pixel, then scanning the whole image, changing the marking of pixel which has already been labeled as the big marking value into small marking value, the marking counter subtracts one.

Step4: All pixels carry out the 2nd step .when all pixels processing are completed, the algorithm is over.

2) Object feature and shape judgment condition setting

After connected component detection algorithm, we should scan the whole marking image to count the area ,barycentric coordinates, upper left coordinate and lower right coordinate of rectangle frame which belongs to different connected components with different marking value.

In order to extract real head information, we choose object area and shape characteristics as object feature. If the connected component isn't in line with the shape attribute, then we judge it as false object and the counter subtracts one, else if it is in line with area classification judgment condition, then we judge it as object, else judge it as false object and the counter subtracts one. Allowing that two head possible

connected together, we count the average area of head (avgs2) when the connected component was in line with the shape attribute condition. If avgs2 is in line with the following judgment condition of two people connected together, then we judge the connected component region as two people and the counter adds two.

3) Marking object region

After the above processing, we get the real head object information, including the area, barycentric coordinates, upper left coordinate and lower right coordinate of rectangle frame. Using the coordinates of rectangle frame, we can mark the real rectangle frame region including object features. Image marking module is the foundation of the following people counting module.

D. People Counting Module

From the above analysis we can draw the conclusion that the value of marking counter is the number of people head and can receive the head average area of object through taking the average of the sum about the head pixels in those rectangle frames, therefore the outputs of the system include the image size, the number of people head and the goal (number of people) average area. In people counting module, the number of people head is the people counting result we need. The method for counting people is a robust method and low cost for using a single camera, which can be used in complex indoor spaces.

4. CONCLUSION

Uptill now people counting system is used for the purposes such as to facilitate security management as well as urban planning. In military application for instance in urban warfare, soldiers might not be able to check every room of building. Sending a camera into a room that could autonomously report how many people are present can help soldiers assess threat level.

But apart from this we can use this system in shopping malls. We can count number of people going in particular section and if there is too much crowd in that section then we can segregate the crowd by applying some technique. For example in shopping mall if too much crowd is in ladies clothes especially for jeans clothes, then we can suggest the manager to provide separate section for jeans clothes, so that crowd get segregated. We can also find out which section is more crowded on particular day of a week. Also we can use this system in public places such as cyberbar, computer room, laboratory classroom, conference room etc.

5. REFERENCES

[1] Lin SF, Chen Y, Chao HX, "Estimation of number of people in crowded scenes using perspective transformation,"

IEEE Transactions on Systems, Man and Cybernetics- Part A : Systems and Humans, vol.3l,no 6,2001, pp. 645- 653

[2] T.zhao,R.Nevatia, 'Tracking multiple humans in complex situation," IEEE Transactions on Pattern Analysis and Machine Intelligence , vol .26,no 9, 2004,pp.l208-1221

[3] T.Zhao,R.Nevatia,B,Wu, "Segmentation and tracking of multiple humans in crowded environment," IEEE Transactions on Pattern Analysis and Machine Intelligence,vol .30,no 7, 2007 ,pp.l198-1211

[4] Vijay Mahadevan, Nuno Vasconcelos, "Background subtraction in highly dynamic scenes,"IEEE conference on Computer Vision and Pattern Recognition (CVPR 2008), 2008, pp.I-6

[5] Wei Di, Rongben Wang, "Driver Eyes Identification Based on Infrared Illuminator, "International conference on Computational Intelligence and Software engineering (CISE),2009,pp. 1-4

[6] Matt Wheeler, Michael A. Muda, "Processing colour and complex data using mathematic morphology," National Aerospace and Electronics Conference, 2000, pp. 618-624

[7] Emma Regentova, Shanhram Latifi, Shulan Deng andDongsheng Yao, "An algorithm with reduced operations for connected components detection in ITU-T group 3/4 coded images, "IEEE Transactions on Pattern Analysis and Machine Intelligence, vol .24,no 8,2002,pp. 1039-1047

[8] K. Terada, D. Yoshida, S. Oe, and J. Yamaguchi, "A method of counting the passing people by using the stereo images," Proceedings of the International Conference on Image Processing,1999. ICIP 99., vol. 2, pp. 338–342 vol.2, 1999.

[9] S. Velipasalar, Y.-L. Tian, and A. Hampapur, "Automatic counting of interacting people by using a single uncalibrated camera," IEEE International Conference on Multimedia and Expo, 2006, pp. 1265–1268, July 2006.

[10] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003., vol. 2, pp. II–459–66 vol.2, June 2003.

[11] T. Zhao and R. Nevatia, "Tracking multiple humans in crowded environment," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004., vol. 2, pp. II–406–II–413 Vol.2, June-2 July 2004.

[12] V. Rabaud and S. Belongie, "Counting crowded moving objects," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, vol. 1, pp. 705–711, June 2006.

[13] D. Kong, D. Gray, and H. Tao, "Counting pedestrians in crowds using viewpoint invariant

training," in 18th International Conference on Pattern Recognition, ICPR, pp. 1187– 1190, 2006.

[14] P. A. Mehta and T. J. Stonham, "A system for counting people in video images using neural networks to identify the background scene," Journal of Pattern Recognition, vol. 29, no. 8, pp. 1421–1428, 1996.

[14] T. Schl¨ogl, B. Wachmann, H. Bischof, and W. Kropatsch, "People counting in complex scenarios," Technical report, pp. 1–8, 2003.

[15] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," Proceedings of the

International Conference on Computer Vision and Pattern Recognition, pp. 1–7, 2008.

[16] J. W. Kim, K. S. Choi, B. D. Choi, and S. J. Ko, "Realtime vision-based people counting system for the security door," Proceedings of International Technical Conference On Circuits Systems Computers and Communications, 2002.

[17] V. Rabaud and S. Belongie, "Counting crowded moving objects," Proceedings of the International Conference on Computer Vision and Pattern Recognition, pp. 705 – 711, 2006.

[18] S. Harasse, L. Bonnaud, and M. Desvignes, "People counting in transport vehicles," Proceedings of the International Conference on Pattern Recognition and Computer Vision, pp. 221–224, 2005.

[19] D. Tsishkou, L. Chen, and E.Bovbel, "Semi-automatic face segmentation for face detection in video," International Conference on Intelligent Access to Multimedia Documents on the Internet, pp. 107–118, 2004.

Mr. Akhil Khare is working as an Associate Professor in Information Technology Department at Bharati Vidyapeeth Deemed University College of Engineering, Dhankawadi, Pune India. He was awarded his Master of Technology Degree from RGTU Bhopal. He is persuing his PhD from JNU,Jodhpur. His areas of interest are Computer Network, Software Engineering and Multimedia System. He has nine years experience in teaching and research. He has published more than twenty research papers in journals and conferences. He has also guided ten postgraduate students.

Kanchan Warke from Information Technology Department at Bharati Vidyapeeth Deemed University College of Engineering, Dhankawadi, Pune India.. Her areas of interest are Software Engineering and Multimedia System. She has Five years experience in teaching and research.

Dr. Akhilesh R. Upadhyay obtained Ph.D. degree from the Swami Ramanand Teerth Marathwada niversity, Nanded in 2009, M.E. (Hons.) and B.E. (Hons.) in Electronics Engineering from S.G.G.S. Institute of Engineering & Technology, Nanded [M.S.] in year 2004 and 1996 respectively. He is currently working as Vice Principal and Head of Electronics and Communication Engineering Department at SIRT, Bhopal, India. He has more than 12 years teaching and 3 years of industry experience. He is Associate Editor of Journal of Engineering, Management & Pharmaceutical Sciences, Ex-Editor of International Journal of Computing Science and Communication Technologies and member of editorial boards/review committee of various reputed journals and International conferences. He has more than 50 research publications in various international/national journals and conferences; he also authored more than 16 text/reference books on electronics devices, instrumentation and power electronics. He is recognized Ph.D. Supervisor for various Universities in India and presently guiding 11 Ph.D. scholars.

Anisotropic Dry Etching (RIE) for Micro and Nanogap Fabrication

Th. S. Dhahi[1], U. Hashim[2]

[1,2] Institute of Nano Electronic Engineering, University Malaysia Perlis (UniMAP)

ABSTRACT

The main objective of this research is to develop a micro and nanogap structure using dry anisotropic etching –Reactive Ion Etching- RIE. Amorphous silicon material is used in the micro and nanogap structure and gold as electrode. The fabrication processes of the micro and nanostructure are based on conventional photolithography, wet etching for the Al pattern and wet etching for a-Si pattern using RIE process. Reactive ion etching (IP-RIE) has been applied and developed as essential method for etching micro and nanogap semiconductors. The fabrication and preparation methods to fabricate micro and nanogaps using RIE properties are discussed along with their advantages towards the nanotechnology and biodetection. In this research, 2 masks designs are proposed. First mask is the lateral micro and nanogap and the second mask is for gold pad electrode pattern. Lateral micro and nanogaps are introduced in the fabrication process using amorphous silicon and gold as an electrode. As a result we need to deposit Al layer over the amorphous silicon semiconductor material before coating a photoresist to protect the a-Si layer during the etching and using the Al layer as a hard mask. The requirement time to etch 1μm amorphous silicon pattern completely by using IP-RIE to fabricate the micro and nanogap structure its take approximately 30sec. These results are better than those using wet anisotropic etching techniques.

Keywords **- Micro and nanogap, photolithography, reactive ion etching (RIE).**

1. INTRODUCTION

The RIE equipment used in these experiments was a Vacutech parallel-plate system. The lower electrode is powered by a 13.56 MHz-RF generator coupled through an automatic tuning network. Each electrode is 200 mm in diameter and the distance between them is 23 mm. The RF electrodes are made of anodised aluminium. The chamber volume is 131 and the system is pumped by a 350 l/rain lurbomolecular pump backed by a mechanical rotary vane pump. The base pressure before each run was less than 5 x 10 Tort. A two-level factorial design of experiments was used to find the main and interaction effects governing etch rate and sidewall slope. Even through the RIE process performance is influenced by interaction effects between different factors, such as, oxygen content, power density, pressure and loading, at a chosen set of etching parameters.

Different techniques and schemes such as, sandblasting, mechanical grooving, and wet chemical etch; laser sculpturing and plasma etching have been used to texture the surface of single and polycrystalline silicon. One of the characteristics of all these methods is that they use either chemical- or physical-etching mechanisms to produce the desired pattern. This imposes limitations such as features geometry or selectivity of the process depending on which mechanism is chosen [1].

In recent years, nano-ordered materials have attracted much attention since the products have been smaller in keeping with a trend of density growth or integration in various technology fields. They have great interest because of their potential to exhibit novel properties which cannot be achieved by bulk materials. For examples, two-(2D) and three-dimensionally (3D) ordered materials have attracted much interest due to their potential applications in photonic crystals [2–5], data storage [6–9], field emission device [10–13]. It is important to establish fabrication techniques of materials with desired shape or size depending on each application. For example, they must have periodic pillar or hole structures with high aspect ratio to use as 2D photonic crystals [11, 12]. It is necessary to be moth eye structure for use as antireflection coating [13, 14], which have pointed top preferably for field emission device [15].

Micromachining in a-silicon and silicon are widely employed for the fabrication of various micromechanical structures needed for many types of sensors and actuators Pattern transfer based on various etching processes plays an essential role in micromachining in the development of etching

processes for micromachining it is important to maintain high etch rate, good control of line width, good uniformity and high selectivity over both mask and underlying layer However, most of the traditional wet etchants are unable to meet these requirements for several reasons First, there is pure chemical reaction involved in the wet etching, resulting in an isotropic etching profile unless a crystal orientation dependent etch is used Secondly, adhesion of photoresist to the substrate is often poor due to the attack of the etchant Thirdly, surface tension of the liquid makes it impossible for the wet etchant to penetrate through very small windows in resist pattern and react with substrate Fourthly, gas bubble formation during the etching locally prevents the etching from proceeding and leads to poor uniformity Although some of the wet etchants for bulk micromachining, such as KOH or EDP, etch crystal silicon anisotropically, their etching characteristics strongly depend on the crystal orientation, doping concentration and electrical potential of the substrate As a result, the type, shape and size of the structures that can be realized are limited [16].

In previous reports, a variety of techniques for narrow nanogap fabrication have been demonstrated: electron beam lithography [17], [18], electromigration [19], mechanical break junction [20], sacrificial layer-assisted silicon and gold nanogaps [21], and surface-catalyzed chemical deposition [22]. However, except for electron beam lithography and sacrificial layer-assisted nanogaps, all other techniques have several problems in nanogap commercialization because of the complex steps and difficulties in fabricating reproducible nanogaps and their compatibility with other semiconductor circuits and processes. Therefore, new approaches and integration [23] methods for fabricating nanogap arrays need to be developed in order to overcome these problems.

On the other hand metallic nanoparticles have been used to establish self-assembly nanostructure of which physical and chemical properties have been investigated in recent years. In particular, gold nanoparticles can be easily prepared and have the characteristic of biocompatibility. Some new devices have been developed for the application of immunoassay by making use of the novel properties of nanostructure that is self-assembled by gold nanoparticles [24]. However, the properties of gold nanostructure can vary significantly with the size of gold nanoparticles and the pitch between gold nanoparticles in the nanostructures [24–26]. A gold-amplified sandwich immunoassay for the detection of human immunoglobulin G has also been developed by Natan and coworkers [26]. The sensitivity of their immunoassay can reach 1 pM. They have also conducted a series of studies on the effect of particle size and surface coverage on the detective sensitivity of immunoassay [27, 28].

A gap in a material with a high magnetic permeability causes magnetic field lines to leak into the surrounding ambient environment (air). If this gap is varied in size, the density of the field at a given point in that area will change. MEMS fabrication techniques enable the design and fabrication of a device that can take advantage of this behavior and locally modulate a static magnetic field [29].

2. METHODOLOGY

In this research, Si wafer is used to fabricate a micro and nanogap biosensor. The first step is to design and produce a mask, which is two mask designs are proposed, the amorphous silicon micro and nanogap with gold electrode process flow are develops. This research is mainly focus on the issue related to the fabrication of the micro and nanogap and the development of a new technology. The sidewall etching using IP-RIE to form thin micro and nanogap metal cantilevers which configured the 3-D micro and nanogap electrode grid array structure. Anisotropy of RIE is modeled and the etching profiles are simulated.

The starting material used in this research is a P-type, 100 mm in diameter (4 inch wafer) silicon-on-insulator (SOI) wafer as shown in Fig. 1.
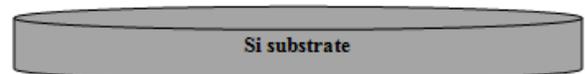
Figure 1: Si wafer

Silicon substrate (Si) wafer is used to reduce parasitic device capacitance and thus improve the final device performance.

The first process is to check the wafer type from its specification, measure wafer thickness (Si thickness), measure the sheet resistance. After that, lightly scribe the backside of each wafer; protect the top surface, using the scribe tool provided. Mark gently but make it visible and place scribed wafer in container.Wafer is cleaned before each process.

As for the lithography process, two photomasks are employed to fabricate the micro and nanogap using conventional photolithography and a-Si dry etching techniques. Commerical chrome mask is be used in

this research for better photomasking process. This mask is used to develop the gold electrode with a-Si micro and nanogap pattern. The photomasks are designed using AutoCAD and then printed onto a chrome glass surface.

Fig. 2 is the first mask for micro and nanogap electrode formation which the length and width of 5000μm and 2500μm respectively.
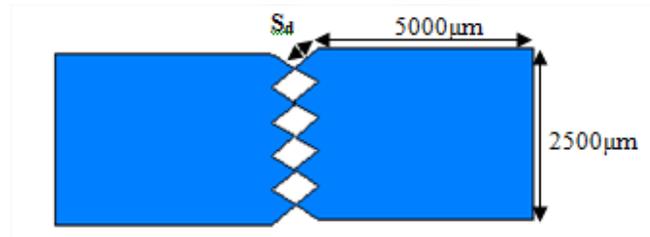
Figure 2: Design specification of the first Mask.

The proposed angle length of the end electrode are shown in Table 1. This is simply to check the best angle for the best micro and nanogap formation after etching process.

Table 1: Difference dimensions for S_d.

S_d	1	2	3	4	5	6
μm	600	700	800	900	1000	1100

The symbol S_d refer to the dimension for side angle of the design for micro and nanogap formation. Its show that when S_d is large this mean the micro and nanogap become very sharp and less sharp with less dimension of S_d.

Fig. 3 shows the actual arrangement of device design on chrome mask. It is consist of 160 dies with 6 different designs.
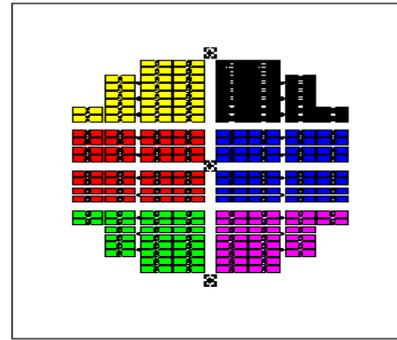
Figure 3: Schematic nanogap design of the actual mask1on chrome glass

Fig. 4 is a schematic device design of mask 2 with 5000μm length and 2500μm width. The distant between two rectangles is indicated as S_a bearing the same dimension with S_d according to the theory of Pythagoras, and the dimension of S_a can be defined mathematically as shown in figure 5.
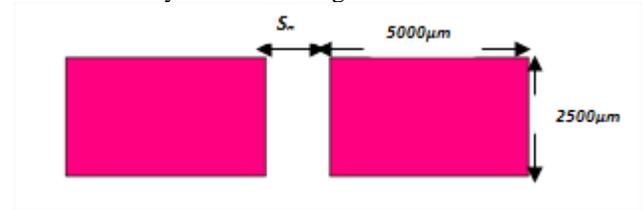
Figure 4: Design Specification for Mask2
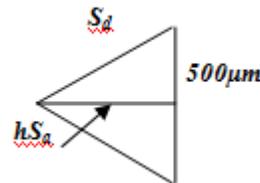
Figure 5: Schematic representation S_a, where $S_a=2hS_a$

Table 2: Variance Dimensions for S_a and S_d

S_d (μm)	600	700	800	900	1000	1100
$hS_a=((S_d)^2-(250)^2)^{1/2}$ (μm)	545	653	759	864	968	1071
$S_a=2 hS_a$ (μm)	1090	1307	1516	1729	1963	2139

From the above table the dimension for S_a depend on the dimension of S_d. The calculated S_a is based on $S_a=2hS_a$

Fig. 6 is a schematic mask on chrome glass.
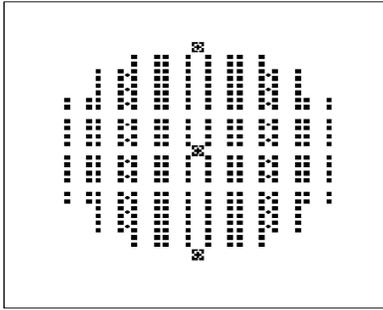
Figure 6: Schematic electrode Mask 2 on chrome glass

3. RESULTS AND DISCUSSIONS

The proposed process steps of gold electrode with a-Si micro and nanogap fabrication are shown in Figure 10. After cleaning the Si wafer, deposit 100nm silicon oxide layer over the Si wafer by using PECVD equipment then deposit 50nm a-Si layer on the silicon oxide surface using the same equipment PECVD before applied photolithography process, a layer of positive photoresist is first coating the a-Si surface, and then exposed to ultraviolet light through a mask 1. After development only the unexposed resist will remain. After that, applied dry etching process for a-Si pattern to fabricate the micro and nanogap for the micro and nanostructure using the recipe's parameters as explain in the Table 3.

Table3: RIE for a-Si recipe.

Cf_4	0
CHF_3	0
SF_6	50
O_2	10
Ar	30
Bais	250
Power ICP Power	650
APC/Control (Pa)	1.20
Etching Time	10 Sec

After remove the resist we can show by using SEM some damage in the photoresist layer for the pattern as shown in Fig. 7.
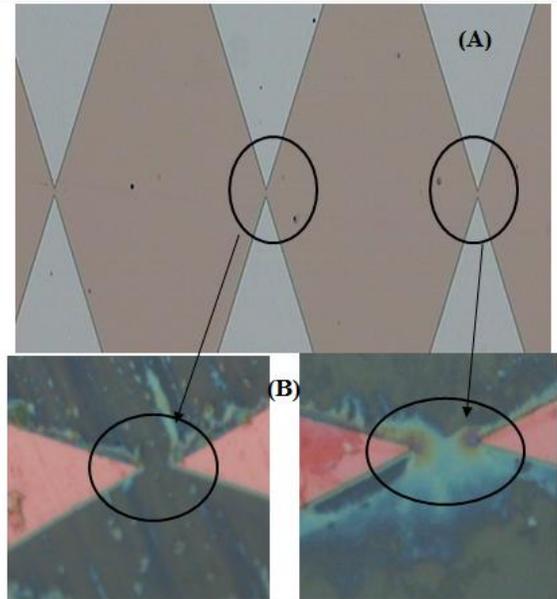
Figure 7: SEM photo for a-Si micro and nanogap:
A) Before using RIE process, B) after using RIE process.

It is clearly show that there is a problem in the gap and the size of the gap pattern, this affects negatively on the results of the examination for the electrical device characterization, and the proposed is depositing 135nm Al layer as a hard mask to avoid damage of the photoresist layer and the micro and nanogap structure during etching process by using the RIE. Next, in the photolithography process, a layer of positive photoresist is first applied on to the Al surface, and then exposed to ultraviolet light through a mask 1. After development only the unexposed resist will remain. After that a wet etching process of AL layer is performed using Al etch solution before removing the resist. After that, applied dry etching process for a-Si pattern by using same the recipe's parameters in the Table 3 to fabricate the micro and nanogap for the micro and nano structure, then applied wet etching to remove the Al layer, and Figure 8 show SEM image after applied dry RIE for the a-Si layer.
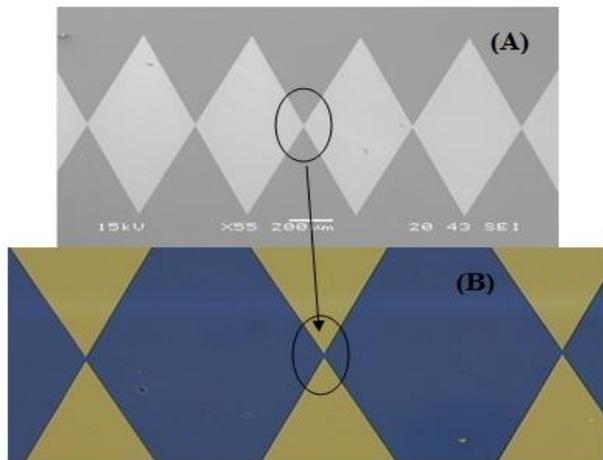
Figure 8: SEM image for a-Si micro and nangap pattern: (A) before using dry etching RIE, B) after using RIE.

It is clearly show that the change that took place in the amorphous silicon design after the deposition of the Al layer as a hard mask and how it has become sharper, and of course these positive results in the etching process lead to good results in the biodetection process because the RIE is an essential process in the micro and nano fabrication. Fig. 9 shows the planned the amount of fish and supply which was etching in the RIE process.
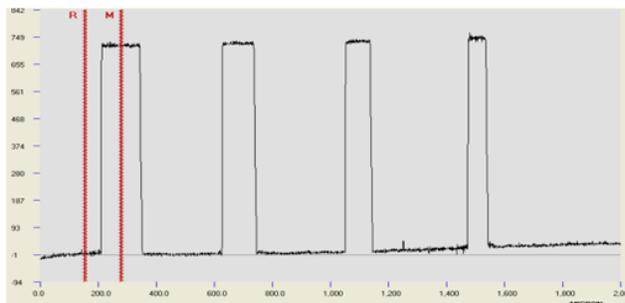

Figure 9: the thickness measurements for a-Si pattern after dry etching process.

The figure above it is clear that the RIE process applied to more than 70.0nm of the a-Si layer thickness and display the shape engraved up to 2000µm. After fabricate the gap, we proceed to fabricate the gold electrode where deposit a layer of 30nm/100nm for Ti/Au substrate as a first step to design the electrode, following the resist coating process. After exposing mask2 the layer of the resist is developed, then wet etching process of Ti/Au substrate is performed before removing the resist. Finally a

structure of the gold electrode with a-Si micro and nanogap is obtained as in Fig. 10(h).
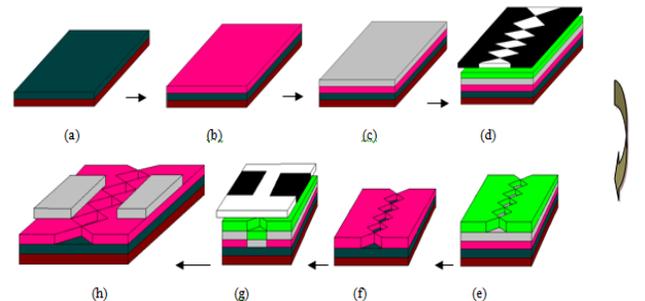
Figure 10: gold/ a-Silicon/ Silicon oxide structure process flow

An important issue in device fabrication is the ability to remove dry etch induced-damage in order to restore the electronic properties of the material. Really, the patterned Al was acted as a hard mask for protect the photoresist layer and a-Si pattern during the etching process. Laterally etch rate is obtained by measuring of image sizes difference between before and after etching with using SEM. In addition, the whole etching processes were done using IP reactive ion etching system with etching parameters of basic recipe as shown in Table 3 which then was further developed to control the nanostructure dimension.

Etch process needs a perfect result of just one step of nano patterns development, since stepped image development can lead to contaminations and then many porosities and un-etched areas result. On the other hand, resist patterns that were developed for sufficiently long time led to resist patterns that were removed of adjusted positions and they were even frequently solved in developer solution. However, optimum development time depends strongly on many parameters such as usage age of developer, resist thickness, soft bake time and density of patterns. In addition, higher patterns density led to longer development and etching times.

From the observation of experiment, the RIE mechanism is assumed to consist of both chemical and physical reaction in the chemical reaction, radicals as active gas molecules react with a-Si molecules and reacted formations are removed, consequently, the process becomes isotropic. Furthermore, this chemical reaction is activated by an impact of ions, that is, a kind of ion sputtering effect. This additive action is physical and has a directional characteristic owing to an incident angle of ions, usually perpendicular to the

electrodes (vertical). Under a certain etching condition, both the chemical and the physical reaction take place simultaneously, therefore, the etched shape may be decided by the ratio of each etch rate that can be defined as an instantaneous etch factor.

4. Conclusion

IPC-RIE has been developed successfully as a method to etch a-Si in a homogeneous way, independently of the grain orientation on the surface. A-Si micro and nanogap structure was produced and fabricated in this work. The etching rate measurements explain the stripping 1μm take 30sec for a-Si layer. These results are better than those using wet anisotropic etch techniques. The fabrication and development of micro and nanogap structure by using reactive ion etching (RIE) process need to deposit Al layer as a hard mask to avoid the damage during the etching.

5. ACKNOWLEDGMENT

REFERENCES

[1] S. Winderbaum, O. Reinhold and F. Yun, Reactive ion etching (RIE) as a method for texturing polycrystalline silicon solar cells, *Solar Energy Materials and Solar Cells 46,* 1997, 239-248.

[2] E. Yablonovitch, Inhibited Spontaneous Emission in Solid-State Physics and Electronics, *Phys. Rev. Lett. 58,* 1987, 2059.

[3] V.V. Poborchii, Photonic-band-gap properties of two-dimensional lattices of Si nanopillars, *J. Appl. Phys. 91,* 2002, 3299.

[4] F. Pommereau, L. Legouezigou, G.H. Duan and B. Lombardet, Fabrication of low loss two-dimensional InP photonic crystals by inductively coupled plasma etching, *J. Appl. Phys. 95,* 2004, 2242.

[5] H.C. Guo, D. Nau, A. Ranke and H. Giessen., Large-area metallic photonic crystal fabrication with interference lithography and dry etching, *Appl. Phys. B. 81,* 2005, 271.

[6] I. Soten, H. Miguez, S.M. Yang, S. Petrov, N. Coombs, N. Tetreault, N. Matsuura, H.E. Ruda and G.A. Ozin, Barium Titanate Inverted Opals - Synthesis, Characterization, and Optical Properties, *Adv. Funct. Mater. 12,* 2002, 71.

[7] T. Aoki and M. Kuwabara, Micropatterned epitaxial (Pb,La)(Zr,Ti)O$_3$ thin films on Nb-doped SrTiO$_3$ substrates by a chemical solution deposition process with resist molds, *Appl. Phys. Lett. 27, 2004,* 2580.

[8] T. Aoki, M. Kondo, K. Kurihara, N. Kamehara and M. Kuwabara, Crystallinity of Microscopically Patterned (Pb,La)(Zr,Ti)O$_3$ Films on (001)Nb-Doped SrTiO$_3$ Substrates Prepared by Chemical Solution Deposition Process with Resist Molds, *J. Appl. Phys. 45,* 2006.

[9] J.E. Jang, S.N. Cha, Y. Choi, G.A.J. Ameratunga and D.J. Kang, Nanoscale capacitors based on metal-insulator-carbon nanotube-metal structures, *Appl. Phys. Lett. 87,* 2005, 263103.

[10] A.G. Rinzler, J.H. Hafner, P. Nikolaev, L. Lou and R.E. Smalley, Unraveling nanotubes: Field emission from an atomic wire, *Science 269,* 1995, 1550.

[11] V. Mizeikis, S. Juodkazis, J.Y. Ye, A. Rode, S. Matsuo and H. Misawa, Silicon surface processing techniques for micro-systems fabrication, *Thin Solid Films 438,* 2003, 445.

[12] V.V. Poborchii, A visible–near infrared range photonic crystal made up of Si nanopillars, *Appl. Phys. Lett. 75,* 1999, 3276.

[13] H. Toyota, K. Takahara, M. Okano, T. Yotsuya and H. Kikuta, Fabrication of Microcone Array for Antireflection Structured Surface Using Metal Dotted Pattern, *J. Appl. Phys. 40,* 2001, 747.

[14] D.H. Macdonald, A. Cuevas, M.J. Kerr, C. Samundsett and D. Ruby, Texturing industrial multicrystalline silicon solar cells, *Solar Energ.76,* 2004, 277.

[15] Minoru Mizuhata, Takuya Miyake, Yuki Nomoto and Shigehito Deki, Deep reactive ion etching (Deep-RIE) process for fabrication of ordered structural metal oxide thin films by the liquid phase infiltration method, *Microelectronic Engineering, 85,* 2008, 355–364.

[16] Y.X. Li, M.R. Wolffenbuttel, P.J French, M. Laros, P.M. Sarro and R.F. Wolffenbuttel, Reactive ion etching (RIE) techniques for micromachining applications, *Sensors and Actuators A, 41-42,* 1994, 317-323.

[17] P. B. Fischer and S. Y. Chou, 10 nm Electron Beam Lithography and Sub-50 nm Overlay Using a Modified Scanning Electron Microscope, *Appl. Phys. Lett., Vol. 62, No. 23,* 1993, 2989-2991.

[18] S. Itousa et al., Fabrication and AFM Characterization of a Coplanar Tunnel Junction with a Less Than 30 nm Interelectrode Gap, *Nanotechnology, vol. 5,* 1994, 19-25.

[19] H. Park et al., Fabrication of Metallic Electrodes with Nanometer Separation by Electromigration, *Appl. Phys. Lett., vol. 75, No. 2,* 1999, 301-303.

[20] M. A. Reed, Zhou C, Muller CJ, Conductance of a molecular junction, *Science, vol. 278 (5336),* 1997, 252-254.

[21] C.W. Park et al., Fabrication of Poly-Si/Au Nano-gaps Using Atomic-Layer-Deposited Al2O3 as a Sacrificial Layer, *Nanotechnology, 16*, 2005, 361-364.

[22] C.S. Ah et al., Fabrication of Integrated Nanogap Electrodes by Surface-Catalyzed Chemical Deposition, *Appl. Phys. Lett, 88,* 2006, 133116-1–113116-3.

[23] J. Kim et al., Integration of 5-V CMOS and High-Voltage Devices for Display Drive Applications, *ETRI Journal, vol. 20, 1*, 1998, 37-45.

[24] P.E. Buckle, R.J. Davies, T. Kinning, D. Yeung, P.R. Edwards, D. Pollard-Knight and C.R. Lowe. The resonant mirror: a novel optical sensor for direct sensing of biomolecular interactions part II: applications, *Biosens. Bioelectron. 8,* 1993, 355.

[25] K. Choi, H.J. Youn, Y.C. Ha, K.J. Kim and J.D. Choi, Detection of cholera cells using surface plasmon resonance sensor, *J. Microbiol. 36,* 1998, 43.

[26] L.A. Lyon, M.D. Musick and M.J. Natan, Colloidal Au-Enhanced Surface Plasmon Resonance Immunosensing, *Anal. Chem. 70*, 1998, 5177.

[27] L.A. Lyon, M.D. Musick, P.C. Smith, B.D. Reiss, D.J. Pena and M.J. Natan, Surface plasmon resonance of colloidal Au modified gold films, *Sensor. Actuator. B 54,* 1999, 118.

[28] L.A. Lyon, D.J. Pena and M.J. Natan, Surface Plasmon Resonance of Au Colloid-Modified Au Films: Particle Size Dependence, *J. Phys. Chem. B 103,* 1999, 5826.

[29] C.D. White, G. Piazza, P.J. Stephanou and A.P. Pisano, Design of nanogap piezoelectric resonators for mechanical RF magnetic field modulation, *Sensors and Actuators A 134,* 2007, 239–244.

REFERENCES

This heading is not assigned a number.

A reference list **MUST** be included using the following information as a guide. Only *cited* text references are included. Each reference is referred to in the text by a number enclosed in a square bracket (i.e., [3]). References **must be numbered and ordered according to where they are first mentioned in the paper,** NOT alphabetically.

Examples follow:

Journal Papers:

[1] M Ozaki, Y. Adachi, Y. Iwahori, and N. Ishii, Application of fuzzy theory to writer recognition of Chinese characters, *International Journal of Modelling and Simulation, 18(2),* 1998, 112-116.
Note that the journal title, volume number and issue number are set in italics.

Books:

[2] R.E. Moore, *Interval analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1966).
Note that the title of the book is in lower case letters and italicized. There is no comma following the title. Place of publication and publisher are given.

Chapters in Books:

[3] P.O. Bishop, Neurophysiology of binocular vision, in J.Houseman (Ed.), *Handbook of physiology,* 4 (New York: Springer-Verlag, 1970) 342-366.
Note that the place of publication, publisher, and year of publication are enclosed in brackets. Editor of book is listed before book title.

Theses:

[4] D.S. Chan, *Theory and implementation of multidimensional discrete systems for signal processing*, doctoral diss., Massachusetts Institute of Technology, Cambridge, MA, 1978.
Note that thesis title is set in italics and the university that granted the degree is listed along with location information

Proceedings Papers:

[5] W.J. Book, Modelling design and control of flexible manipulator arms: A tutorial review, *Proc. 29th IEEE Conf. on Decision and Control*, San Francisco, CA, 1990, 500-506.
Note that the proceedings title is set in italic

"Role of Vital Factors for the Success of Products of Small Entrepreneurs"

B.S.Ravikiran[1], Dr.C.B.Vijaya Vittal[2]

[1]Research Scholar, Dept of Mechanical Engg,SSIT,Tumkur-5,Karnataka
[2] Professor & Dean (R&D), HMSIT, Tumkur, Karnataka

ABSTRACT:

Small Entrepreneurs are major contributors to the economic growth and job creation. In this research an attempt is made to explore the factors and strategies contributing to the success and failures of the products of small entrepreneurs. It does not identify any industry specific strategies and factors, managerial abilities or other specific characteristics related to successful operation of small entrepreneurs. This paper provides guidelines for the success of the products for small entrepreneurs. This could help to improve the ability of small entrepreneurs to develop and prosper in an increasing competitive and complex world. A model has been developed to forecast the success or failure of the product which will be useful for small entrepreneurs.

Keywords: *Entrepreneur, forecasting model success factors,*

1. INTRODUCTION:

The end result of a manufacturing process is a product to be offered to the marketplace to satisfy a need or want. Thousands of new products are introduced to the market every year. Many small entrepreneurs developing new products and modification to the existing products has become a necessity and a way of life. Discovering which factors or practices lead to business success and failure is a primary and yet unfilled purpose of business. Understanding user needs, external and internal communications, product advantages and marketing efforts have been found to be related to the product success of small entrepreneurs[1].The context was India, a developing nation bound in a multitude of traditions and inertia. In spite of the importance and magnitude of the monetary expense ,the area of new products is still fraught with failures, risks and difficulties [2].Entrepreneur are able to spot options and create new directions for an industry. Typically they deal with ambiguity and change and that is a prerequisite for success in today's fast paced business world. They can distinguish real from imaginary pitfalls and the brightest among them can

turn error into opportunity[3].Entrepreneurs always operate at the edge of their competence , focusing more of their resources and attention on what they do not yet know(eg; investment in R&D) than on controlling what they already know. They measure themselves not by the standards of the past but by visions of the future. Innovation is an essential ingredient for today's social and economic growth. It improves the quality of life, raise standard of living and enables entrepreneur to grow and prosper. Innovation is creating and introducing new ways of doing things, better use of goods, more efficient services and systems. Innovation use knowledge and information. It is desirable to develop a model that enables accurate prediction of the outcome of a new product before heavy expenditures are incurred [4]. Though there are many models to predict the success of the products of big Entrepreneurs all existing models require large number of data to forecast and hence there is need to have model to visualize the products at the idea stage itself based on the innovators thinking and theirs capabilities with single set of data. An attempt has been made to predict the success of the products of small entrepreneur based on a single data.

2.Research Methodology:

This research relied on primary data collected by the survey method. The data was collected from the users about the product of small entrepreneurs.The first survey data were collected from users of arecanut peeling machine. A set of 52 questionnaires was prepared. These questionnaires were grouped into eight factors viz; Consumer, Government Role, Economics of the product, Physical characteristics, Attributes of the product, marketing of the product, Entrepreneur's attribute, Environmental condition. Consumer factors refer to the consumer's purchasing capacity of the product, status of the consumer. Government role refers to certifications and support from the Government. Economics of the product refers to the cost resale value, fuel consumption savings in time. Physical characteristics refer to

weight, compactness, space occupation, availability in different size and quantity. Attributes of the product refers to reliability, robustness, safety, efficacy, adaptability, repairability. Marketing of the product refers to after sales service, resale value, self repairable. Entrepreneur's attribute refers to the investment capacity of the Entrepreneur, his capability to take risk, his capability of involvement etc. Environmental condition refers to labor availability, Government policies. A five point Likert scale ranging from 1=Unsatisfactory to 5=excellent was used to measure the extent to which users respond to each variable. The users were from different locations, varying economic condition and rural background. The users were personally contacted and interviewed. They were given the set of questionnaire and asked to fill up the questionnaire and their opinion about the product. The factors are given below:

Sl No	Factors
G1	Consumer
G2	Government Role
G3	Economics of the product
G4	Physical Characteristics
G5	Attributes of the product
G6	Marketing of the product
G7	Entrepreneur's attribute
G8	Environment condition

Addresses of users of the products were obtained from the entrepreneurs who manufacture the product and market on their own. Arecanut peeling machine was taken for the research purpose. The small entrepreneurs are V-tech Thirthahalli, Dharma Technologies, Tumkur, SR Agrotech, Tumkur .These entrepreneur's machine was approved by Agriculture Department,Govt of Karnataka. They have produced innovative products namely Arecanut peeler, Arecanut polisher,Mini tipper.

3.Results and Discussion:
3.1. Reliability of the data:
Using Reliability calculator the reliability and validity of the data was found.The Cronbach alpha

was found out to be 0.9545.This means that the data collected was reliable and valid.

3.2. Correlation Coefficient:
The correlation Coefficient analysis was carried out. The Pearson Correlation Coefficient between the groups was obtained. It was found that G4 & G3,G5 & G3, G6& G5,G8 & G5 are strongly correlated as the Pearson Coefficient is greater than 0.7 .

3.3. Regression Analysis:
The Regression analysis was done to predict the success of the product. Considering G7 (Entrepreneur's attribute) as dependent variable and other variables as independent variable a multiple regression model was obtained in the form of an equation:

$$G7 = 8.44 - 0.027\ G1 + 0.221\ G2 + 0.232\ G3 - 0.168\ G4 - 0.0503\ G5 + 0.111\ G6 + 0.042\ G8$$

Predictor	Coef	SE Coef	T	P
Constant	8.440	3.934	2.15	0.055
G1	-0.0273	0.1585	-0.17	0.867
G2	0.2211	0.1879	1.18	0.264
G3	0.2323	0.1270	1.83	0.095
G4	-0.1678	0.1250	-1.34	0.207
G5	-0.05025	0.08421	-0.60	0.563
G6	0.1114	0.2385	0.47	0.649
G8	0.0420	0.1681	0.25	0.807

$S = 1.16415$ R-Sq = 59.1%   R-Sq(adj) = 33.0%

PRESS = 38.3453   R-Sq(pred) = 0.00%

3.4. Hypothesis Testing:
This test was conducted between two entrepreneurs who are leading manufacturers of arecanut peeling machine namely V-Tech and Dharma Technologies. The test was conducted on 3 Groups of factors namely;

G1	Role of consumer
G4	Physical Characteristics

G6	Marketing of the product

The hypothesis are :

H1: There is no significant difference between two companies with respect to the role of Consumer

H2: There is no significant difference between two companies with respect to the physical characteristics of the products

H3: There is no significant difference between two companies with respect to the marketing of the products

a.G1-Consumer

Two-Sample T-Test and CI: g1, g1_1

Two-sample T for g1 vs g1_1

	N	Mean	StDev	SE Mean
g1	7	22.57	3.74	1.4
g1_1	7	24.14	2.34	0.88

Difference = mu (g1) - mu (g1_1)

Estimate for difference: -1.57143

95% CI for difference: (-5.28348, 2.14062)

T-Test of difference = 0 (vs not =): T-Value = -0.94 P-Value = 0.368 DF = 10

H1:There is no significant difference between the two entrepreneurs with respect to the role of consumer as P value is >0.1

b.G4-Physical Characteristics

Two-Sample T-Test and CI: g4, g4_1

Two-sample T for g4 vs g4_1

	N	Mean	StDev	SE Mean
g4	7	38.29	4.19	1.6
g4_1	7	36.86	2.91	1.1

Difference = mu (g4) - mu (g4_1)

Estimate for difference: 1.42857

90% CI for difference: (-2.06769, 4.92483)

T-Test of difference = 0 (vs not =): T-Value = 0.74 P-Value = 0.476 DF = 10

H2: There is no significant difference between the two entrepreneurs with respect to physical characteristics of the product as P value is >0.1

C.G6-Marketing of the Product

Two-Sample T-Test and CI: g6, g6_1

Two-sample T for g6 vs g6_1

	N	Mean	StDev	SE Mean
g6	7	13.14	2.12	0.80
g6_1	7	13.86	2.19	0.83

Difference = mu (g6) - mu (g6_1)

Estimate for difference: -0.714286

90% CI for difference: (-2.782698, 1.354127)

T-Test of difference = 0 (vs not =): T-Value = -0.62 P-Value = 0.548 DF = 11

H3: There is no significant difference between the two entrepreneurs with respect to marketing of the product as P value is >0.1

3.5. Analysis of Variances (ANOVA):
One way ANOVA analysis was done with respect to dependent variable G7 and an independent variable G2

One-way ANOVA: G7 versus G2

Source	DF	SS	MS	F	P
G2	6	22.92	3.82	3.40	0.034
Error	12	13.50	1.13		
Total	18	36.42			

S = 1.061   R-Sq = 62.93%   R-Sq(adj) = 44.40%

The relevant graphs are shown below:

Residual Plots for G7
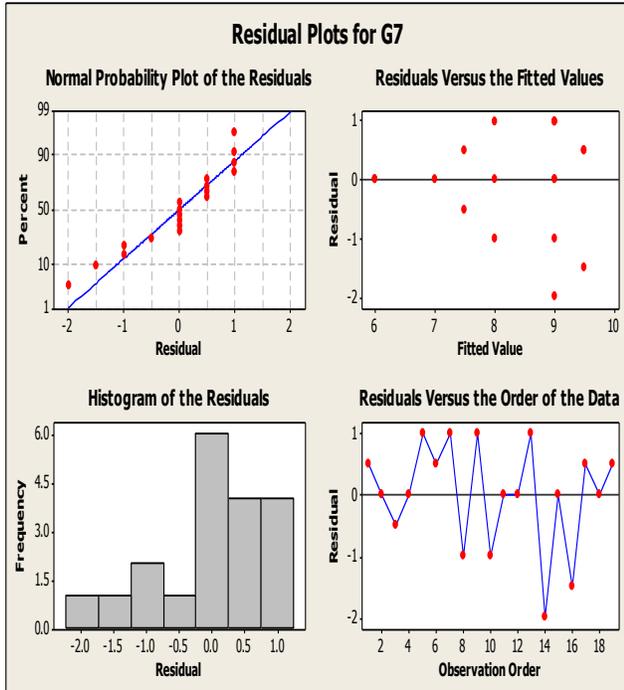
It is found that G2 and G7 are related to each other as p value is near to zero. But the other independent variables are not closely related like G2 which is as shown below

One-way ANOVA: G7 versus G5

Source	DF	SS	MS	F	P
G5	13	23.42	1.80	0.69	0.726
Error	5	13.00	2.60		
Total	18	36.42			

S = 1.612   R-Sq = 64.31%   R-Sq(adj) = 0.00%

One-way ANOVA: G7 versus G8

Source	DF	SS	MS	F	P
G8	7	18.50	2.64	1.62	0.227
Error	11	17.92	1.63		
Total	18	36.42			

S = 1.276   R-Sq = 50.81%   R-Sq(adj) = 19.50%

One-way ANOVA: G7 versus G1

Source	DF	SS	MS	F	P
G1	7	16.33	2.33	1.28	0.344
Error	11	20.10	1.83		
Total	18	36.42			

S = 1.352   R-Sq = 44.83%   R-Sq(adj) = 9.71%

4. CONCLUSION:

It is found that for the success of product Entrepreneur should concentrate on all eight factors. Each factor has an impact on the success of a product. Especially for a new Entrepreneur Government support is most important. An Entrepreneur should have sufficient resources to convert customer needs to customer demand. The products which have failed lacked in providing the perceived superior advantages or the entrepreneur failed to effectively communicate to the user superior advantages. Entrepreneur lacked the credibility, competence and financial resources. Each of the entrepreneur failed to anticipate the problems in the turnaround of money and the consequence with respect to the successful commercialization of the product. It may be concluded that the entrepreneur should give equal importance to all factors. If he neglects one factor it will have cascading effect on other factors.

REFERENCES:
[1] Radiah Abdul Kadeer,Mohd Rosli Bin Mohamad,Ab.Azid Hi.Che Ibrahim,"Success factors for small Rural entrepreneurs under the one district one industry programme in Malaysia".

[2] Sharad Sarin,Gour.M.Kapur,1990, "Lessons from New Product Failures;Five case Studies", Journal of International Marketing Management,Vol 19, pp 310-313

[3] A Dale Timpe,2005,Book on Creativity,Jaico Publishing House,ISBN 81-7224-863-6

[4] Lo,Frang Cheong Wah,Foo Say Wei,Bauly Jonh.A,"Multiple regression models for electronic product success prediction",IEEE International Conference on Management of Innovation and Technology 200,pp 419-422

[5] Nishika Patel,"Mastermind India",April 2010,India Today,pp 47-90

[6] Arne L.Kalleberg, Kevin T.Leight,1991,"Gender and Organisational performance: Determinants of small business survival and success", Academy of Management Journal,
Vol 34,No1,136-161

[7] Michael P.Steiner and Olaf Solem,January 1988,"Factors for success in small manufacturing firms",Journal of small business management,pp 51-56

[8] Roger Calantone & Robert G.Cooper,1981,"New product scenario: Prospects for success",Journal of Marketing,Vol 45,pp 48-60

[9] Charles Jabani Mambula, 2004, Realting External Support,business Growth & Creating strategies for survival: A comparative case study,Analyses of small Manufacturing firms and Entrepreneur,Small Business Economics,Vol 22,pp 83-109

[10] Attahir Yusuf ,"Critical success factors for small business: perceptions of South pacific Entrepreneurs",April 1995,Journal of small business management,pp 68-73

[11] R.G.Cooper,1979,"The dimensions of industrial new product success and failure",Journal of marketing,Vol 43,pp 93-103

[12] Rosalind C.Paige and Mary A.Littrell,2002, "Craft Retailers' criteria for success and associated business strategies",Journal of small business management,Vol:40(4),pp 314-331.

Reaction Pathways in High Temperature Combustion of Iso-octane

C.R.Berlin Selva Rex[1], A.Haiter lenin[1] ,C.V Manoj Kumar[2],Dr.K.Thyagarajan[3]

1. Research Scholar, Anna University, 2. LBS Institute of Technology for Women, Trivandrum.
3. Research Superviser,Anna University,Tamil Nadu.

Abstract-
This article presents the reaction pathways in high temperature combustion of wide range of hydrocarbon fuel Iso-octane. The activation energies of iso-octane are higher than n-octane. Hence iso-octane is widely used for combustion simulations. For this study a chemical kinetic scheme of iso-octane with 994 elementary reactions and 201 species has been developed and validated with LLNL (Lawrence Livermore National Laboratory) detailed mechanism with 3606 reactions and 857 species. A detailed study on the oxidation and soot formation has been conducted analytically using the reduced chemical mechanism with 994 reactions and 201 species. Species like CH, C2H, C2H2, C3H3, C3H4, C3H6, and C4H6 play a major role in the formation of soot as their decomposition leads to the production of radicals involved in the formation of Polycyclic Aromatic Hydrocarbon (PAH) and the further growth of soot particles. Temperature, pressure, fuel, O2 and OH concentration are also considered in soot formation process. It is also depending upon the nature of air fuel mixture (lean, stoichiometric or rich). A program has been developed in MATLAB for the calculation and prediction of the concentration of 201 intermediate species and the ignition delay in the combustion of Iso-octane. The various initial conditions considered was in between the temperatures of 600K to 1250K with pressure ranging from 10atm to 40atm at various equivalence ratios of 0.3 and 0.6. Nitrogen is considered as the diluent. The diluent percentage is assumed as 79% to make a comparison with atmospheric condition. The criteria for determination of ignition delay times are based on the OH concentrations to reach to a value of 1×10^{-9} moles/cc. The ignition delay times are obtained by varying initial conditions of the mixture in the combustion of Iso-octane. The results on ignition delays have been found to be agreeable with those available in the literature. Cantera (an object oriented software for reacting flows) software is used in this study.

Keywords- *Ignition delay,Reaction mechanism,Iso-octane*

The goal of present study is to establish the understanding of auto ignition in premixed combustion system and the understanding of the behavior of important species which are responsible for soot formation at various initial conditions. Mainly CH, C2H, C2H2, C3H3, C3H4, C3H6 and C4H6 etc are the important species responsible for soot formation. In addition to them temperature, pressure, O2, OH concentration, fuel mixture (lean, stoichiometric, rich) and the fuel structure plays a vital role in soot formation. By increasing the temperature the soot can be burnt out. But the problem is the formation of NOx. By increasing pressure the fuel breaks down in to molecular

level hydrocarbons and radicals and they can easily oxidized, by both OH and O2 in lean mixture and soot formation. Soot is oxidized by OH under fuel-rich and stoichiometric conditions. Soot formation is more at fuel rich zone areas due to the incomplete burning process. Soot is roughly defined as a solid substance that consists of 8 parts of carbon and 1 part of hydrogen. Soot can be formed by 6 process viz. pyrolysis, nucleation, coalescence, surface growth, agglomeration and oxidation. The present study explores oxidation process which converts hydrocarbons to CO, CO2 and H2O.

As computational capacity improves numerical simulations are becoming more attractive for combustion studies. Comprehensive detailed kinetic mechanisms have been compiled to fully describe the fundamental chemical processes involved in fuel oxidation .For example, Curran et al. [1,2] have developed comprehensive mechanisms to study the oxidation of n-heptane and iso-octane. The former mechanism comprised of 560 species and 2539 reactions, while the latter contains 857 species and 3606 reactions. These mechanisms were tested by comparing computed results with various experimental data from laboratory devices, and a reasonably good agreement was reported between the predicted and the measured results, implying that the reaction mechanisms represent correctly the imported reaction pathways and rates of oxidation for these fuels.

With reaction mechanisms that consist of several hundred species and several thousand reactions, it is still much too costly to use a detailed chemical kinetic mechanism directly in engine combustion studies using multidimensional CFD codes. It is necessary to develop chemical reaction mechanisms that retain the essential features of the fuel chemistry predicted by comprehensive reaction mechanisms, but with much improved computational efficiency in terms of memory usage and CPU time. The extent of comprehensiveness depends on the available computational resources and the type of information desired from the simulation.

The basic aim of mechanism reduction is to identify unimportant species and reaction pathways in order to reduce the complexity of chemistry of mechanisms and important features of full schemes. Various methods have been suggested to determine the importance of species and reaction pathways in a mechanism.

Sensitivity analysis [3-4] is one of the earliest methods, which is simple to apply, but requires extensive post processing to provide decoupled information about the reactions and species. A reaction elimination method [5] was suggested to identity optimal sets of reactions under given constraints. But the optimization approach is asymptotically

slower than sensitivity analysis. The method of detailed reduction [6] uses direct comparisons of reaction rates with preselected critical values in order to speed up the identification of unimportant reactions. However this method is likely to neglect important slower reactions that involve crucial radicals. Another technique to reduce the complexity of a chemical reaction mechanism is chemical lumping [7] which simplifies a mechanism by replacing a set of lumped pseudo species. This method is effective for reducing the number of species and still maintains the important features of full schemes when the importance of a single isomeric species is low but the reaction path via all isomers is significant.

Automation of reduction procedures has drawn much interest as well. Soyhan et al.[8] developed an automatic reduction technique to reduce a natural gas mechanism from 53 species and 589 reactions to 23 species and 20 global reactions. Montgomery et al. [9] used the CARM (computer aided reduction method) to automate the mechanism reduction process to generate a variety of reduced ethylene and n-heptane mechanism. More recently, the direct relation graph method [10] has been suggested and used to help automation of reduction procedures.

I. ISO-OCTANE

Iso-octane (2,2,4-trimethylpentane), a primary reference fuel for octane rating in spark ignition engines, has drawn considerable interest as a model compound for branched alkane components found particularly in gasoline [11], but also those found in diesel [12] and jet fuels [13]. Due to its relevance to practical liquid fuels, iso-octane has been the subject of many experimental and kinetic modeling studies. Experimental investigations of iso-octane oxidation and ignition have been carried out in shock tubes, rapid compression machine (RCMs), flames, jet stirred reactors and flow reactors.

II. INVESTIGATIONS

In building up a chemical kinetic model, a number of assumptions have to be made. In fact the assumptions are made so that the conditions simulated correspond to the conditions that are in effect in the experimental side.

A. Reaction Mechanism for iso-octane Oxidation

The complete reaction mechanism (994 reactions and 201species) which involve in the combustion of iso-octane is proposed & validated [14] with LLNL (Lawrence Livermore National Laboratory) detailed mechanism. There are 3606 reactions and 857 species. Figure.1. shows the validation of present mechanism with LLNL detailed mechanism.
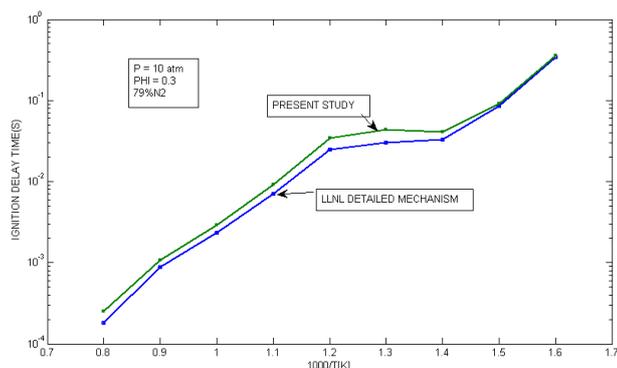

Fig 1: Validation of the reaction mechanism

B. Software

CANTERA object oriented software for reacting flows is used to simulate the results.

III. CRITERIA FOR FINDING IGNITION DELAY

Ignition delay time(t) is the time corresponding to the maximum rate of reactions between CO and O atoms. Induction period is the time at which the temperature had completed about half its total increase, often defined as the time required for a small(ie 1-5%) temperature or pressure rise. Bowman [15] found that ignition delay time to vary inversely with approximately the first power of the propane concentration, to be only slightly dependant upon oxygen concentration and to decrease with increasing pressure and temperature. It is the time to reach the concentration value of OH to 1×10^{-9} moles/cc .It is the time required to start the decomposition of H_2O_2.Figure.2 shows the criteria used for finding out the induction or ignition delay period.
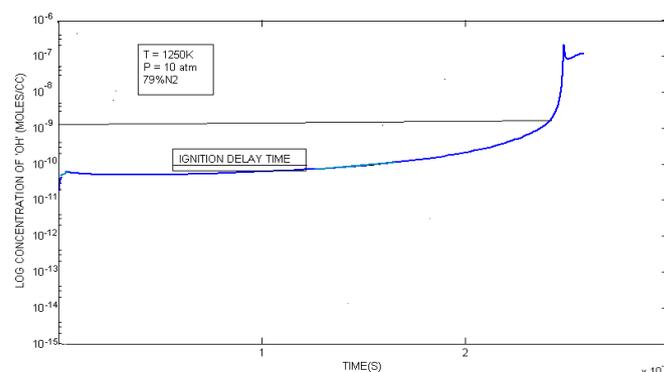

Fig 2: Ignition criteria.

IV. TEMPERATURE PROFILE

The temperature profile is shown in figure.3. The initial condition taken to plot this variation is temperature 1250K,pressure 10 atm and phi=0.3. From the graph it can be seen that the initial dip of temperature occurs. This is due to initiation reaction which are endo-thermic that initially dips and gradually rises until a region when the curve shoots up. These sets of reactions require a large amount of energy and the reactions are endothermic in nature. The heat that is required for the reaction to occur is extracted from the surrounding thus causing the temperature to dip. Once the reaction are complete, a fraction of fuel breaks down in to molecular weight hydrocarbons or the radicals, the temperature steadies on. Once the radicals are formed by the pyrolytic the propagation reaction starts. Initially the reaction proceeds slowly until the self ignition temperature is achieved. Once the temperature is attained, the ignition occurs. The fuel breaks down completely at this stage and the combustion progresses very fast. In the curve, it is the region at which the curve steeps down. The time between the initial time of input and this corresponding time gives the induction period, of course the chemical ignition delay. The curve steeps up and then maintains a particular value. After a while it can be observed that the temperature goes on reducing. This is due to the fact that at this stage there is no fuel to burn to energy. Hence temperature reduces.

The initial condition taken to plot this variation is temperature 1250K,pressure 10 atm and phi=0.3.we can see that at the point of ignition delay, there is a steep increase in temperature and pressure. Sufficient oxygen is not available

to start spontaneous combustion at high equivalence ratio, which increases the ignition delay time.
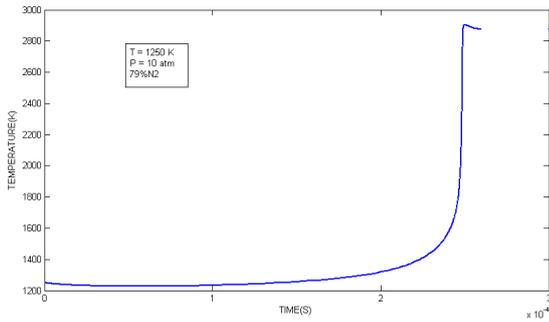
Fig.3. Temperature profile

V. VARIATION OF IGNITION DELAY WITH TEMPERATURE

The variation of ignition delay [16] with temperature is plotted on the figure.8&9 the ignition delay drops when temperature rises. This is so because as the pressure increases, density of fuel oxygen mixtures increases and hence the probability of fuel particle coming in contact with the reacting radicals also increases.

The variation of ignition delay [17] with temperature illustrated in figure.8&9 show that for higher temperature the ignition delay time is decreased. There is profound influence on ignition delay when the initial temperature is increased by 100K.
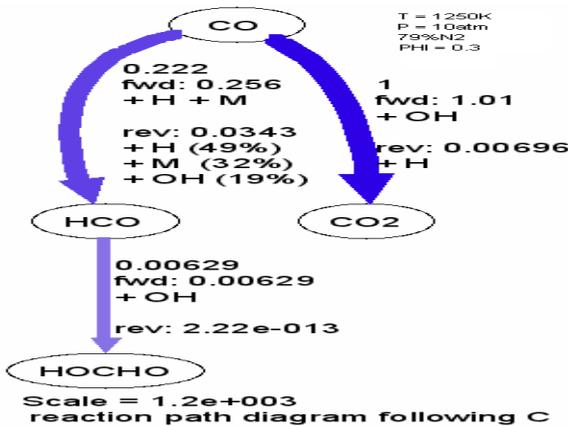
Fig.4.reaction path way of C

Fig.5. reaction path way of H.
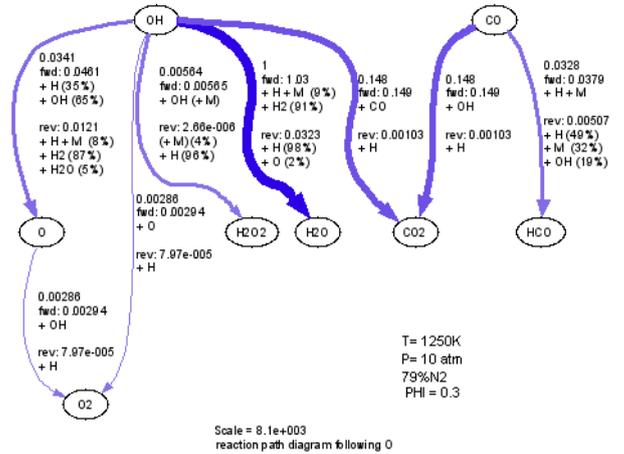
Fig.6.reaction path way of O

VI.SPECIES CONCENTRATION PROFILES

Concentration profiles of major species of C_8H_{18} are show in figure. 7. During the combustion of iso-octane at the point of ignition, the fuel curve steeps down. This proves that ignition has occurred and the fuel is being consumed very fast.

From the figure. 7. the consumption of O_2 is due the combustion of iso-octane, at the point of ignition. The concentration profile of H_2O_2 steeps down because it is also considered as the ignition citeria of iso-octane.
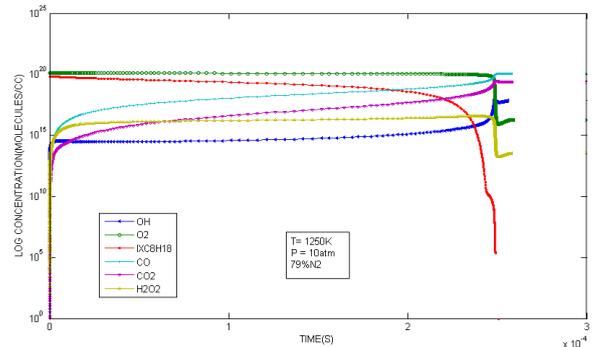
Fig.7 . Species concentration profile
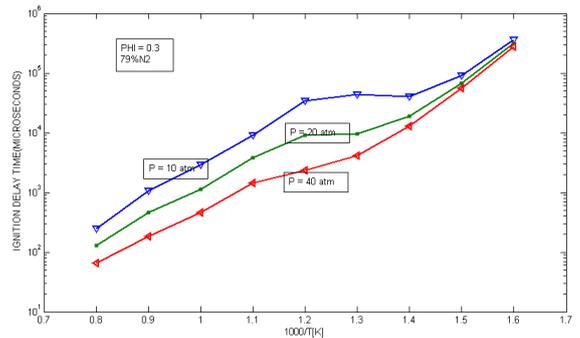
Fig.8 . Variation of ignition delay with temperature

Fig.9. Variation of ignition delay with temperature

Fig.10.variation of ignition delay with equivalence ratio

VII. REACTION PATH WAYS

Figures.4,5&6 shows the reaction path ways of C,H&O Reaction path ways play an important role in determining the unimportant paths which is useful in reaction mechanism reduction.

VIII. SOOT FORMATION

Fig.11.Species Concentration Profile

In the proposed reaction mechanism of 994 elementary reactions and 201 species for Iso-octane the important species which are responsible for formation of soot are plotted in the figure.11, it is observed that the the species C_2H_2 and C_3H_3 are very important in soot formation.

Fig.12. Variation with O2 (lean mixture)
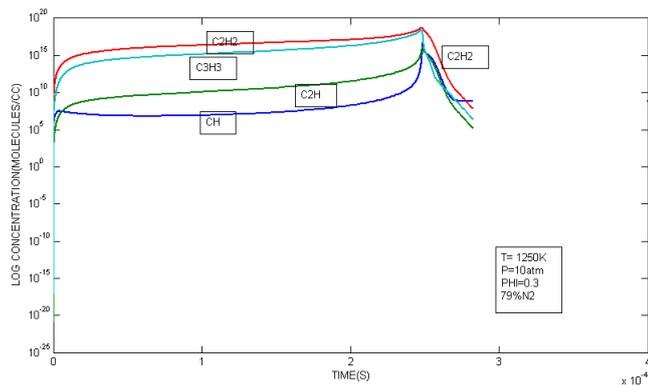
Fig.13. Variation with O2 (Stoichoimetric mixture)
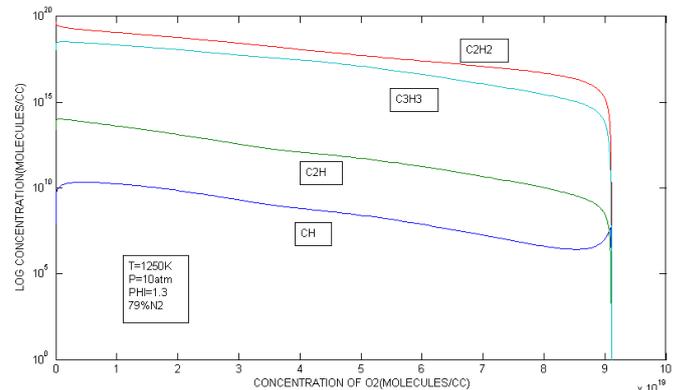
Fig.14. Variation with O2 (Rich mixture)
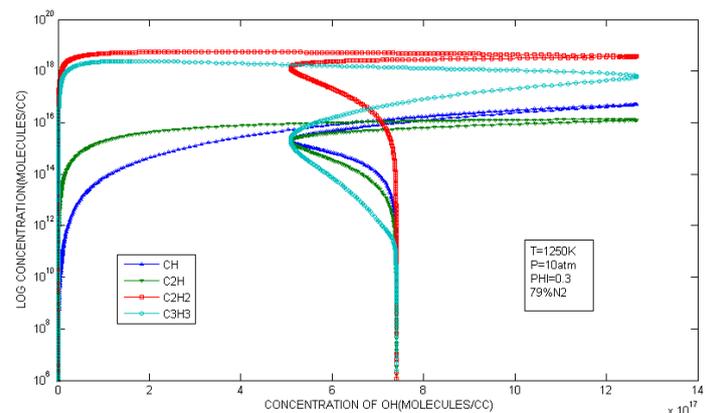
Fig.15. Variation with OH (lean mixture)
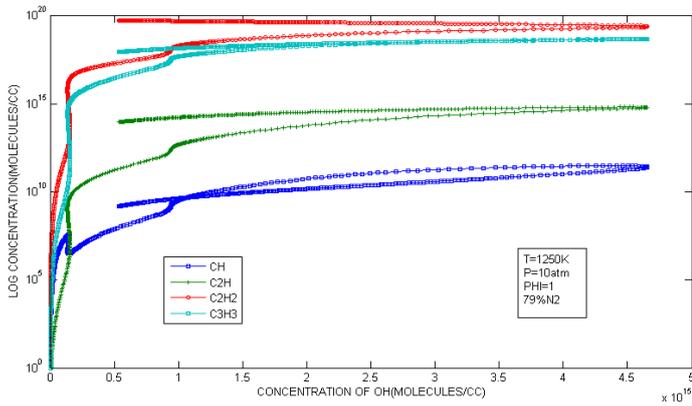
Fig.16. Variation with OH (Stoichoimetric mixture)
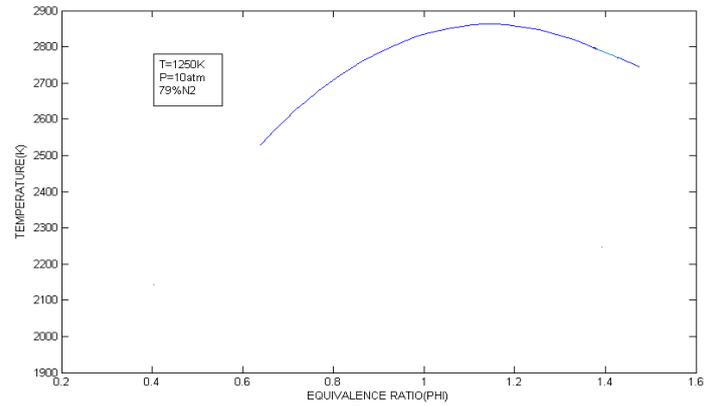
Fig.18.Variation of Temperature with Equivalence ratio
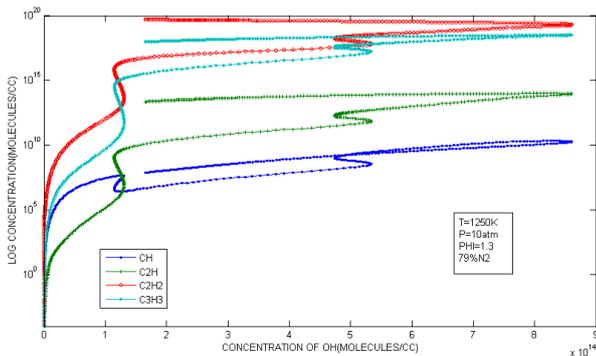
Fig.17. Variation with OH (Rich mixture)

From figures.12,13 and 14 it is observed that concentration level of O2 is maximum at lean mixture because O2 is consumed more for oxidation of soot. At stoichiometric and fuel rich conditions O2 has no role for soot oxidation as shown in figure 13.

At stoichiometric and fuel rich conditions OH radical are responsible for the oxidation of soot is plotted in Figure 16, 17. The concentration levels of OH radical are very high in stoichiomeric and fuel rich conditions as compared with lean mixture condition. Soot is formed by both OH and O2 at lean mixture condition.

Theoretically temperature becomes maximum at equivalence ratio of unity (stoichiometric fuel-air ratio) but using the proposed mechanism it is found that maximum temperature occurs at equivalence ratio of 1.2 as shown in Figure.18. Soot formation increases with equivalence ratio because when equivalence ratio increases fuel percentage in the mixture also increases, if sufficient oxygen is not available for oxidation process unburnt hydrocarbons are produced.

IX. RESULTS

Table 1 : 79%N2, P=10atm, Φ=0.3

Temperature(K)	Ignition Delay Time(µs)
1300	248.6
1100	1065.5
1000	2892.7
900	9093.2
800	3468.97
750	37718.6
700	40563.9
650	92108.4
600	360909.7

Table 2 : 79%N2, P=20atm, Φ=0.3

Temperature(K)	Ignition Delay Time(µs)
1300	130.9
1100	465.3
1000	1121.2
900	3804.9
800	9213.6
750	9585.5
700	18996.5
650	67803.9
600	313171.9

Table 3: 79%N2, P=40atm, Φ=0.3

Temperature(K)	Ignition Delay Time(µs)
1300	65.2
1100	186.3
1000	465.6
900	1452.2
800	1452.2
750	2377.7
700	4096.6
650	13004.9
600	278324.0

Table 4 : 79%N2, P=10atm, Φ=0.6

Temperature(K)	Ignition Delay Time(μs)
1300	305.4
1100	1048.3
1000	2188.8
900	5082.3
800	14904.0
750	18073.7
700	25334.7
650	74836.3
600	314074.4

Table 5 : 79%N2, P=20atm, Φ=0.6

Temperature(K)	Ignition Delay Time(μs)
1300	150.6
1100	426.1
1000	787.9
900	1998.0
800	4107.0
750	5619.9
700	14429.1
650	57178.1
600	270542.5

Table 6 : 79%N2, P=40atm, Φ=0.6

Temperature(K)	Ignition Delay Time(μs)
1300	70.2
1100	160.1
1000	309.7
900	768.7
800	1316.5
750	2869.2
700	10222.3
650	46080.8
600	232320.1

X. CONCLUSION

The analytical study on the reaction pathways in high temperature combustion of iso-octane has been conducted. A reaction mechanism containing 994 reactions among 201 species was proposed and the ignition delay times at various initial conditions of temperatures, pressures and equivalence ratio are determined. The condition used to determine the ignition delay is the time when OH concentration reaches 1×10^{-9} moles/cc .It is found that with increase in initial temperature and pressure during the combustion of Iso-octane, the ignition delay time and formation soot decreases. This is because at high pressure the fuel-air mixture density is increased which result in proper mixing in turn shortens ignition delay times and the formation of soot. With increase in equivalence ratio from lean to rich the ignition delay time and the formation of soot also increases. The reason behind this is when equivalence ratio is increased the fuel concentration in iso-octane is increased and oxygen concentration is reduced so that sufficient oxygen is not available for the combustion to take place quickly. Thus incomplete combustion occurs. In this study the species C2H2 and C3H3 is found to be the most important species for soot formation. A detailed study has been made in this range of temperatures, pressures and the equivalence ratio ranges from 0.3 to 0.6. From the results it is agreeable that for reducing CH emission low

temperature, pressure very lean air fuel mixture is best suited. The results obtained are found to be agreeable with previous studies described herein.

REFERENCES

[1]   H.J. Curran, P. Gauffuri, W.J. Pitz, C.K. Westbrook, Combust. Flame 114 (1998) 149–177.
[2]   H.J. Curran, P. Gauffuri, W.J. Pitz, C.K. Westbrook, Combust. Flame 129 (2002) 253–280.
[3]   T. Turanyi, J. Math. Chem. 5 (1990) 203–248.
[4]   A.S. Tomlin, T. Turanyi, M.J. Pilling, Comprehensive Chemical Kinetics, Elsevier, Amsterdam, 1997, p. 293.
[5]   B. Bhattacharjee, D.A. Schwer, P.I. Barton, W.H. Green, Combust. Flame 135 (2003) 191–208.
[6]   H. Wang, M. Frenklach, Combust. Flame 87 (1991) 365–370.
[7]   S.S. Ahmed, F. Mauß, G. Moréac, T. Zeuch, Phys. Chem. Chem. Phys. 9 (2007) 1107–1126.
[8]   H.S. Soyhan, T. Løvås, F. Mauss, A stochastic simulation of an HCCI engine using an automatically reduced mechanism, ASME 2001-ICE-416, 2001.
[9]   C.J. Montgomery, M.A. Cremer, J.Y. Chen, C.K.Westbrook, L.Q. Maurice, J. Propul. Power 18 (1) (2002) 192–198.
[10] T. Lu, C.K. Law, Combust. Flame 144 (2006) 24–36.
[11] W.J. Pitz, N.P. Cernansky, F.L. Dryer, F.N. Egolfopoulos, J.T. Farrell, D.G. Friend, H. Pitsch, SAE Paper 2007-01-0175, 2007.
[12] J.T. Farrell, N.P. Cernansky, F.L. Dryer, D.G. Friend, C.A. Hergart, C.K. Law, R. McDavid, C.J. Mueller, H.Pitsch, SAE Paper 2007-01-0201, 2007.
[13] M. Colket, J.T. Edwards, S. Williams, N.P. Cernan-sky, D.L. Miller, F.N. Egolfopoulos, P. Lindstedt, K.Seshadri, F.L. Dryer, C.K. Law, D.G. Friend, D.B. Lenhert, H. Pitsch, A. Sarofim, M. Smooke, W. Tsang, AIAA Paper AIAA-2007-0770, 2007.
[14] Hidaka et al-J Mol sci 2,141(1982)
[15] Seery D.J and Bowman, CT'' Ignition delays in several CH4-02-Ar mixtures'', combustion flame, 1970, 14-37.
[16] K A Bhaskaran and P.Roth,The shock tube as wave reactor for Kinetic studies progress in energy and combustion Science,28(2002) 151-192.
[17] E L Peterson,D F Davidson and R K Hanson, Kinetic modeling of shock induced induction in low dilution CH4/02 mixtures at high pressure and intermediate.

Removal of Communication Channel Attack on Biometric Authentication System Using Watermarking

Kamaldeep

Assistant Professor Savera Group of Institutions (Gurgaon)

Abstract — Biometrics authentication system used the physical or behavioral characteristics for identify an individual. With the wide use of biometrics authentication system, the security of biometrics system emerged an important research issue. A biometric system is vulnerable to a variety of attacks aimed at undermining the integrity of the authentication process. In this paper, an approach to enhance the biometrics security by using watermarking technique has been proposed. For hiding the information in the biometrics image taken by sensor, the semi-pixel difference method [1] has been used. In this method, first the pixel is divided into two semi-pixels and their difference is calculated. According to the calculated difference, the watermarked bit is inserted at a pixel value. The pixels for insertion of watermark are selected by using pseudo random number generator which is seeded with a secret key.

After insertion of watermark at sensor module with the help of SPD insertion method it sent to the other end for matching with stored template. The SPD retrieval method finds the watermark. If the watermark is find than the template is original. If the template is tempered by intruder than it is discarded.

Keywords- Biometrics Authentication system, Attacks on biometrcs, Security, Watermarking, Semi-Pixel difference Method etc.

I. INTRODUCTION

A biometrics based authentication system can use physical or behavioral characteristics for identification and verification of a person. It has been deployed in various areas in the industry as well as in military and in the e-commerce. In the current digital world, our biometrics system has a variety of attacks which makes the biometrics system insecure for authentication and communication. With the wide spread utilization of biometrics identification system, establishing the authenticity of biometric data itself has emerged as an important issue. The biometrics system and the attacks on it is shown by the figure 1.
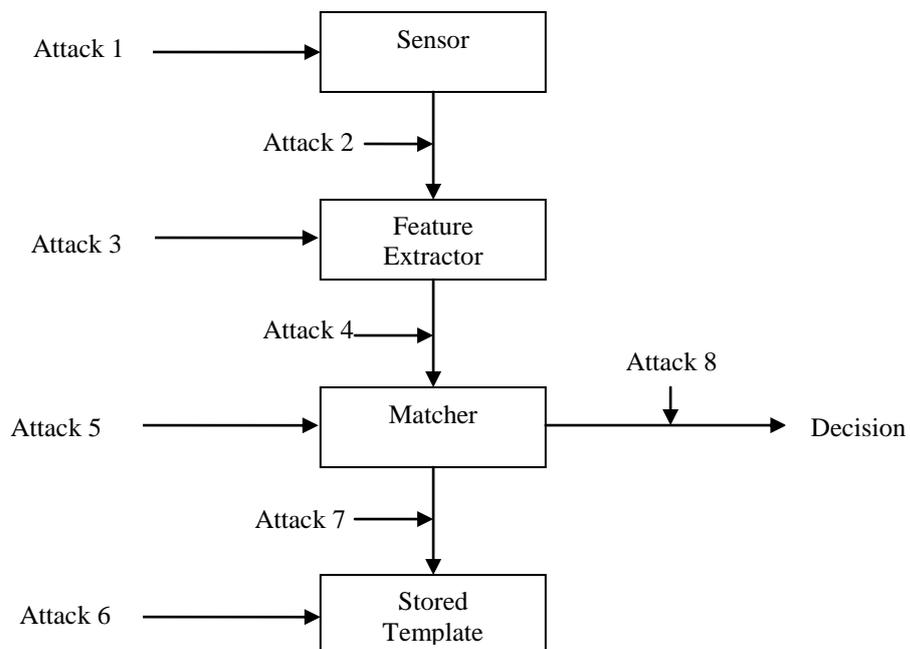
Fig 1. Biometrics authentication system and various attacks on it(Derived from[2])

Digital watermarking or simply watermarking is defined as a process of embedding information like owner name, company logo etc. in the host data. The process of watermark insertion and extraction is given in Figure 2 and Figure 3 respectively [3]. General image watermarking methods can be divided into two groups according to the domain of application of watermarking. In spatial domain methods [4], the pixel values in the image channel(s) are changed. In spectral-transform domain methods, a watermark signal is added to the host image in a transform domain such as the full-frame DCT domain [5]. Watermarking is very similar to steganography in that both seek to hide information in the Cover-object. However steganography is related to secret point-to-point communication between two parties. Thus, steganography techniques are usually having a limited robustness and protect for the embedded information against modifications that may occur during transmission, like format conversion, compression or A/D conversion. On the other hand, watermarking rather than steganography principles is used whenever the media is available to parties who know the existence of the embedded information and may have interest removing it. Thus, watermarking adds additional requirements of robustness. An ideal watermarking system would embed information that could not be removed or altered without making significant perceptual distortion to the media. A popular application of watermarking is to give a proof of correctness of digital data by embedding copyright statements [6].
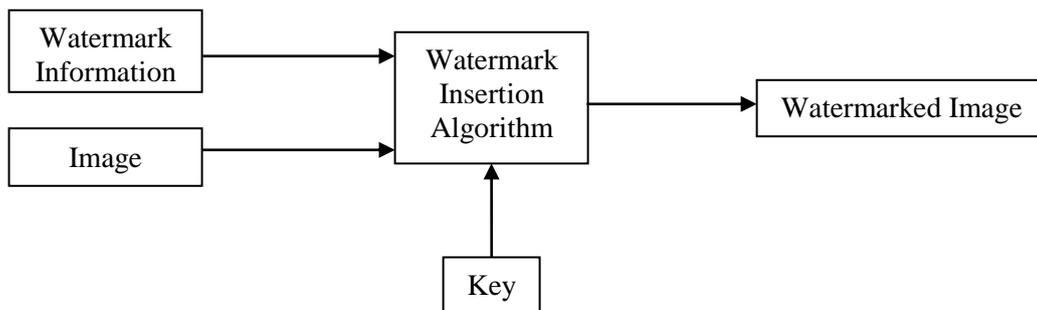
Fig 2. Watermark Insertion Process
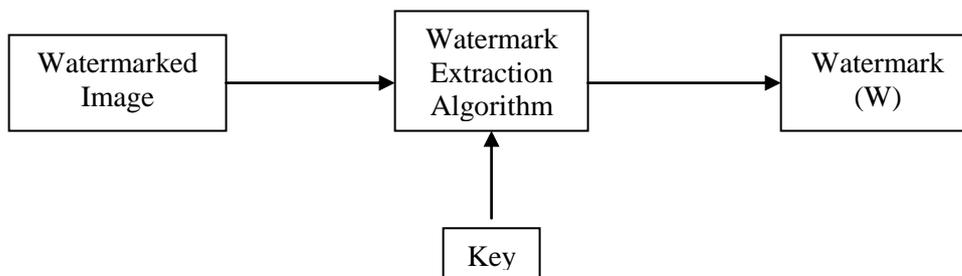
Fig 3. Watermark Extraction Process

The rest of the paper is organized as follows:
     In section 2, use of watermarking in biometrics has been discussed. Section 3 gives Attack no-2 on biometrics system. Section 4 indicates the proposed method. The section 5 discusses the semi-pixel difference method of watermarking. Section 6 discusses the result and conclusion, some emphasis on future work.

II. USE OF WATERMARKING IN BIOMETRICS [7]

     In many cases, the appropriate use of cryptography also reduces this threat [8]. The Security Administrator will configure the biometric system to encrypt and digitally sign all biometric data before it is transmitted from one physical device to another. Steganography can greatly reduce these attacks because attackers must have to obtain the system's private data in addition, to breaching the security of the capture device or biometric storage. This makes these attacks considerably more difficult to achieve But steganography is more secure than cryptography because there is no separate key in steganography rather key is inbuilt in the template [9].
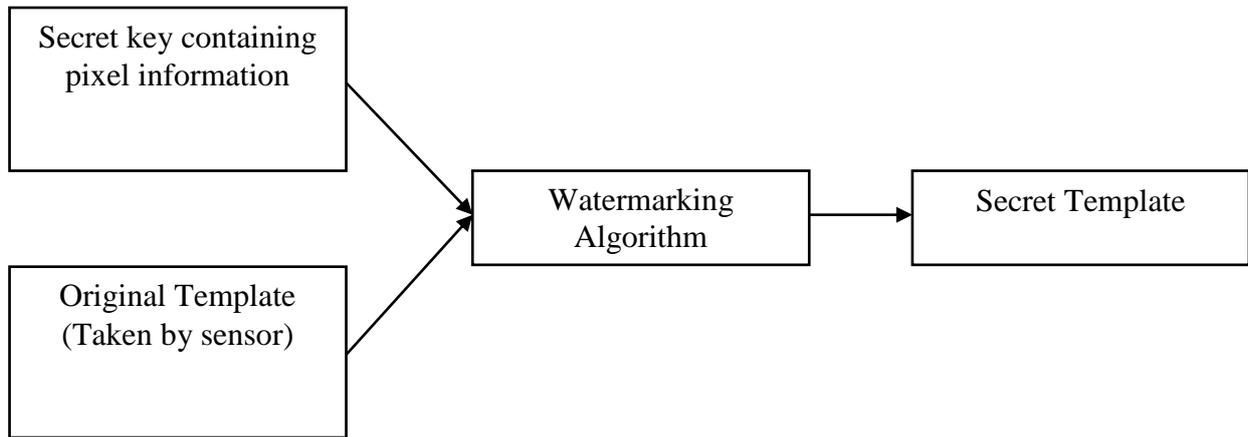
Fig 4. Use of Watermarking in Biometrics

III.   ATTACK NO-2 ON BIOMETRICS SYSTEM.

This point of attack is known as "Attack on the channel between the scanner and the feature extractor" or "Replay attack". In this attack, the attacker intercepts the communication channel between the scanner and the feature extractor to steal biometric traits and store it somewhere. The attacker can then replay the stolen biometric traits to the feature extractor to bypass the scanner. The block diagram of biometrics authentication system and the attack no.2 is given by figure 5.
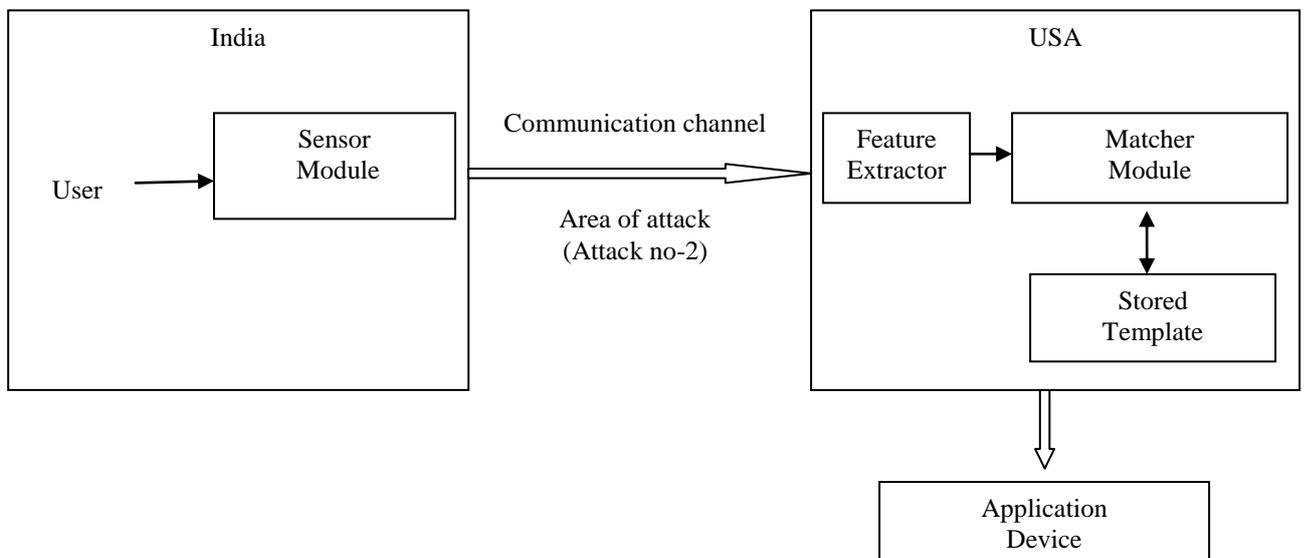
Fig 5. Attack no 2 on the communication channel between sensor module and feature extractor module

IV.   PROPOSED APPROACH

In this section, we gave an approach to remove the communication channel attack i.e. No. 2 Attack on biometric system. Suppose the sensor module is present in INDIA and matching module is present in USA. Then the enrolled template is sent out to USA for matching with stored template. In between intruder can manipulate with the enrolled template. By using watermarking we can remove this problem. The watermark text is inserted within the enrolled template to make it secret template. In the first case, if intruder try to tamper with the secret template then it also changes the watermark text which become visible at the retrieval end i.e. (Matching Module Site.). In, the second case, if intruder replace the original enrolled template then it also become clear at the matching module site because in this case watermark information would be missing from the replaced template. So, by using the watermarking technique we can remove the communication channel attack. The proposed scenario is given by the figure 6.

India

User → Sensor Module

Watermarking Module (Insertion)

Area of attack (Attack no-2)   Communication Channel

USA

Watermarking Module (Extraction and checker)

Feature Extractor

Matcher Module

Stored Template

Application Device
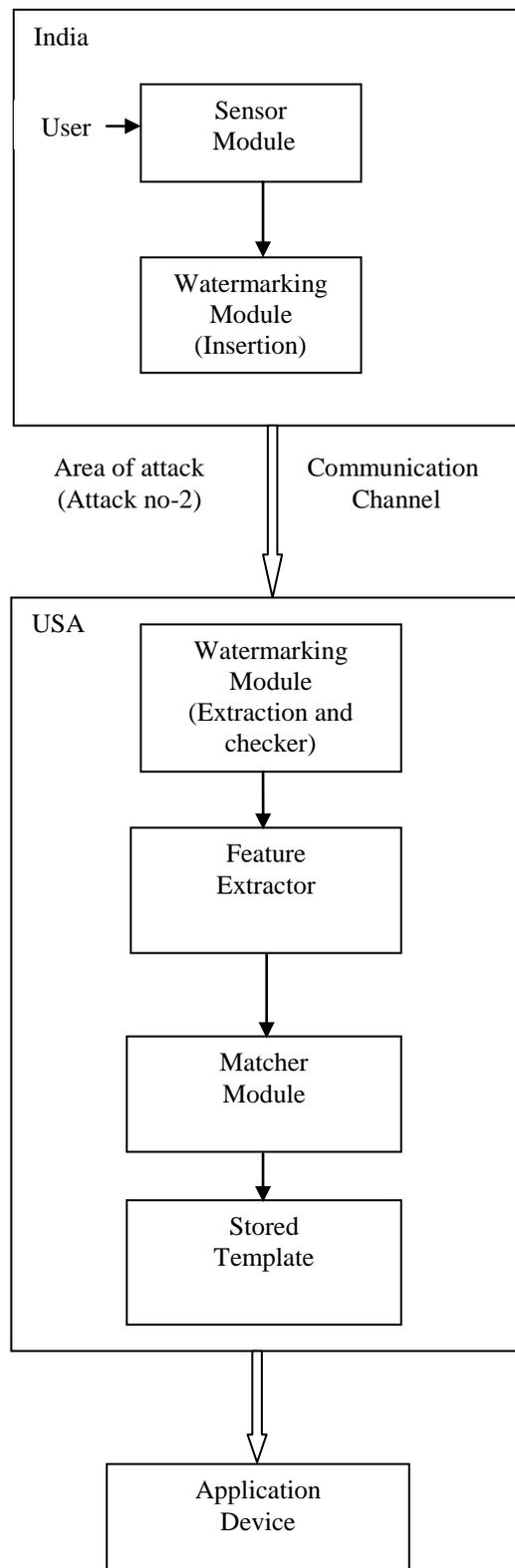
Fig 6. Proposed method for removing the attack no 2 i.e. Attack between sensor and extractor module.

V. SEMI-PIXEL DIFFERENCE METHOD[1]

In this section, watermarking method is described i.e. SPD (Semi-Pixel Difference) method for hiding watermark information in the spatial domain of the gray scale image i.e. biometrics image. SPD method first divides each pixel into two semi pixels known as semi-pixel 1 and semi-pixel 2 and then watermark information is inserted at the pixel value according to the difference of semi-pixel 1 and semi-pixel 2. If we want to insert watermark bit 0 at a pixel value, then the difference of semi-pixel 1 and semi-pixel 2 must be an even number. Otherwise, we made the semi-pixel difference equal to the even number by adding or subtracting 1 to the pixel value. Similarly, if we want to insert watermark bit 1 at a pixel value, then semi-pixel difference must be an odd number otherwise we made the semi pixel difference equal to odd number by adding or subtracting 1 to the pixel value. The pixels for insertion of watermark information are selected by using Pseudo-Random Number Generator that is seeded with a secret key. The split process of pixel is shown in Figure 7. Table I shows how watermark bits can be inserted according to the Semi-Pixel Difference. Figure 8 shows the watermark insertion process & Figure 9 shows the watermark extraction process.
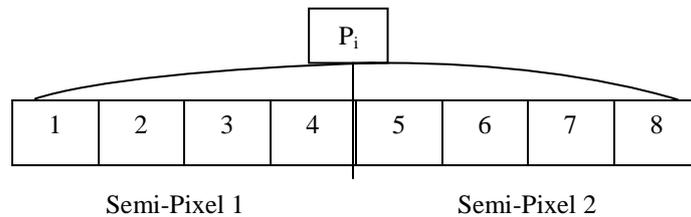
Fig 7. Split Process

Table I: Watermark Insertion according to the Semi-Pixel Difference.

Semi-Pixel Difference	Watermark Bit to be Embedded
0	0
1	1
2	0
3	1
4	0
5	1
6	0
7	1
8	0
9	1
10	0
11	1
12	0
13	1
14	0
15	1

Start

Host Template: H (x, y)
Watermark Text: W_T (b_1, b_2, b_3…..b_n)

for i =1 to n

Get Pixel Location for Watermark Insertion using Pseudo Random Generator

Apply Split Process i.e. Get Semi-Pixel 1 & Semi-Pixel 2.

Calculate SPD=Decimal Value of Semi-Pixel 1 – Decimal Value of Semi-Pixel 2.

Insert bit b_i according to SPD.

i = i+1

Is i = n+1

No

Yes

Watermarked Secret Template: H` (x, y)

Stop
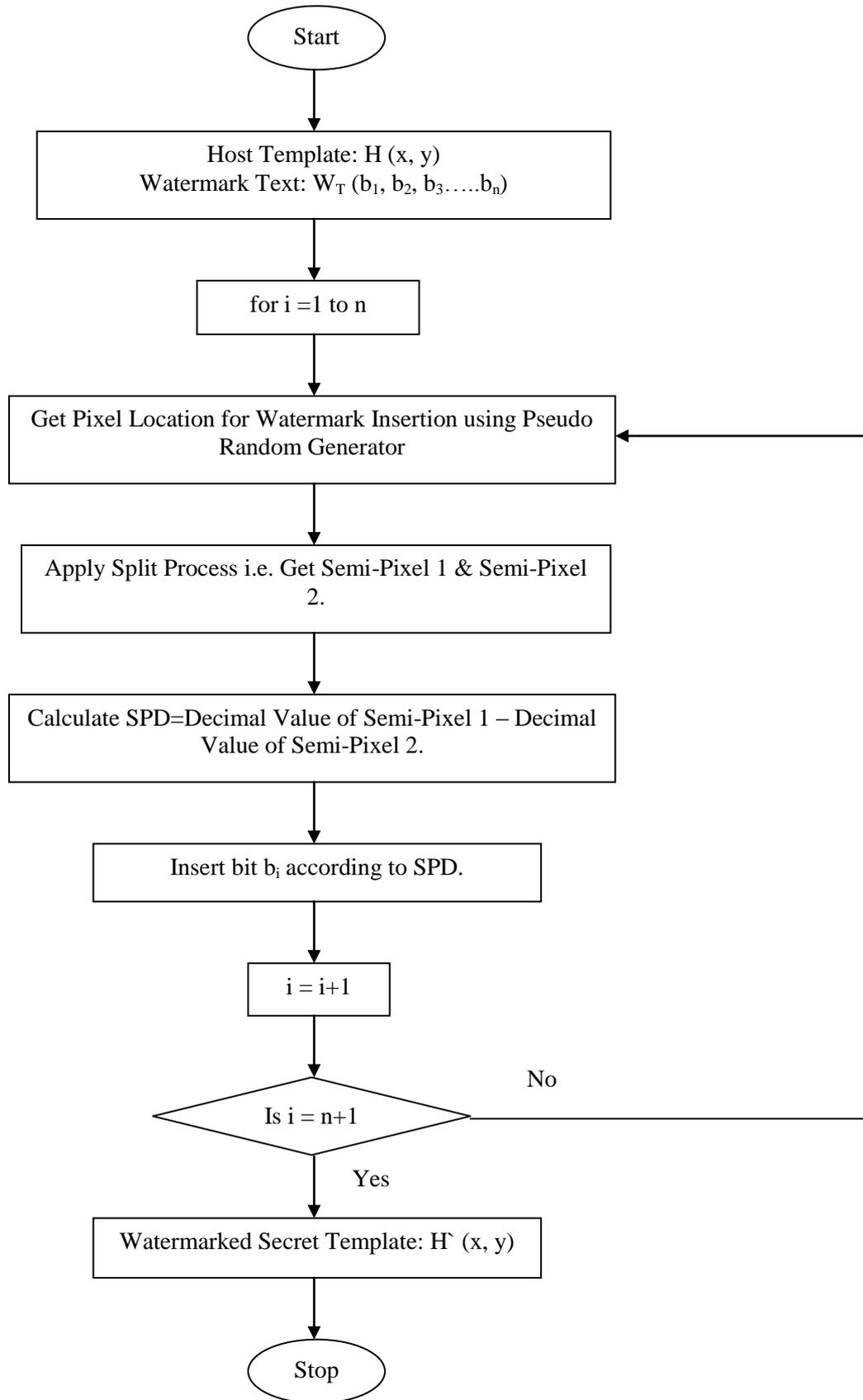
Fig 8. Watermark Insertion Process
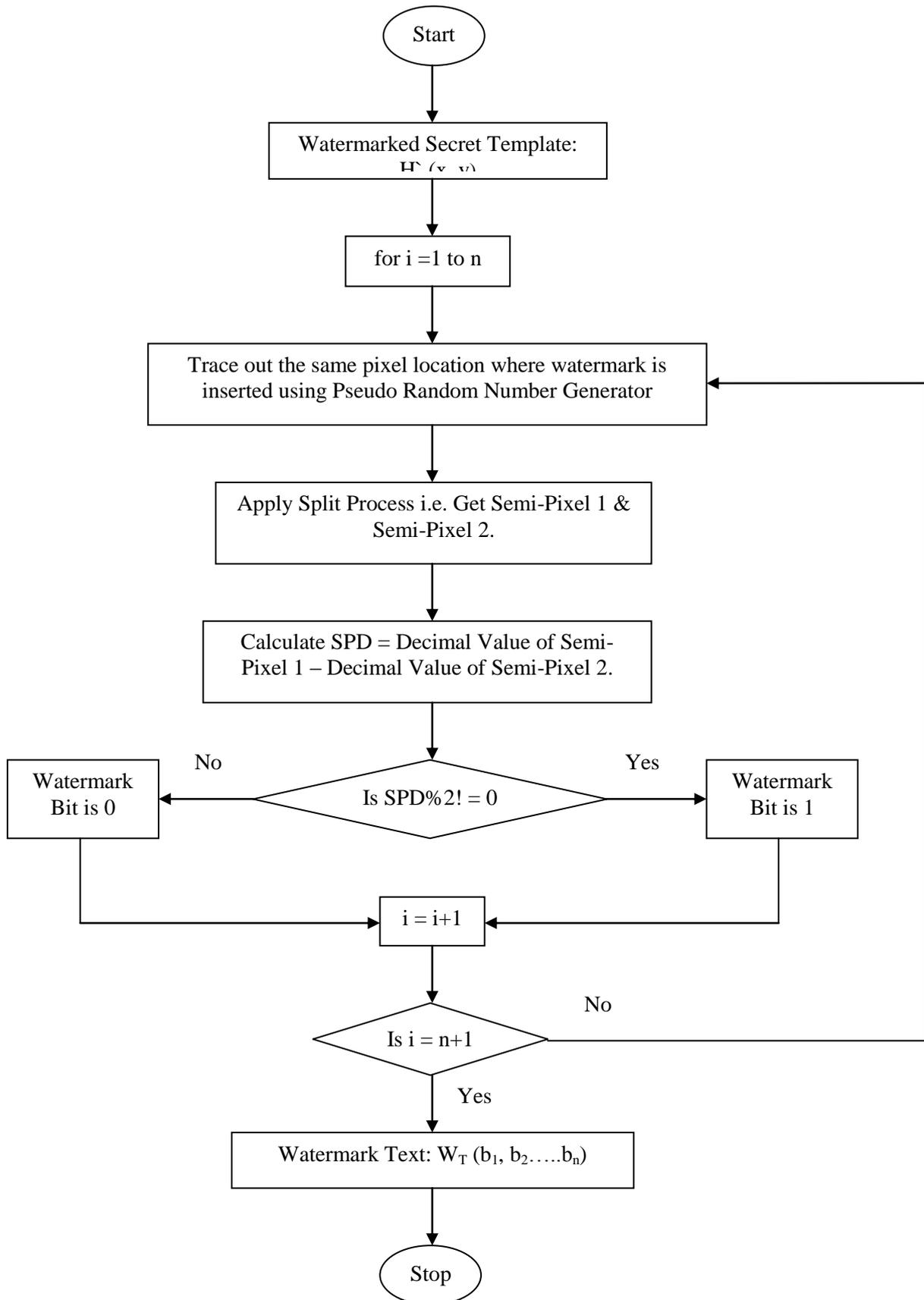
Fig 9. Watermark Extraction Process

5.1 Insertion Algorithm

Step 1: Read the host template: $H(x, y)$

Step 2: Read the watermark text: $W_T(b_1, b_2, \ldots\ldots b_n)$.

Step 3: for i=1to n.

Step 4: Get pixel location P_i for insertion of watermark information using pseudo random number generator.

Step 5: Apply split process i.e. split the pixel into two equal parts i.e. semi –pixel 1 and semi-pixel 2

Step 6: Calculate d_1 and d_2 using equation (1) and (2) respectively.

$$d_1 = DEC \text{ (semi-pixel1)} \text{ ------------ (1)}$$
$$d_2 = DEC \text{ (semi-pixel2)} \text{ ------------- (2)}$$

Step 7: Calculate SPD using equation (3).

$$SPD = Abs (d_1 - d_2) \text{ ------------------- (3)}$$

Step 8: Calculate Decision Variable (DV) using equation (4).

$$DV = SPD \text{ Mod } 2 \text{--------------------- (4)}$$

Step 9: If $b_i=0$ than go to step 10 else go to step 11.

Step10: (a) If DV = = 0, then b_i is present at P_i.

(b) If DV! = 0, then add or subtract 1 to P_i such that DV becomes equal to 0 and insert b_i.

Step11: (a) If DV! = 0 then b_i is present at P_i.

(b) If DV = = 0, then add or subtract 1 to P_i such that DV becomes equal to 0 and insert b_i.

Step 12: Go to step (3)

Step 13: Watermarked secret template: $H`(x, y)$.

Step 14: END.

5.2 Extraction Algorithm

Step 1: Read the watermarked secret template: $H`(x, y)$

Step 2: for i=1to n.

Step 3: Trace out the same pixel location P_i using pseudo random number generator where watermark information is present.

Step 4: Apply split process i.e. split the pixel into two equal part i.e. semi –pixel 1 and semi-pixel 2

Step 5: Calculate d_1 and d_2 using equation (1) and (2) respectively.

Step 6: Calculate SPD using equation (3).

Step 7: Calculate Decision Variable (DV) using equation (4).

Step 8: If DV = = 0, then 0 is the watermark bit else 1 is the watermark bit.

Step 9: Go to step (2).

Step 10: Collect the entire watermark bits to get the watermark text: $W_T(b_1, b_2, \text{------} b_n)$.

Step 11: END

VI. CONCLUSION AND FUTURE SCOPE

In this paper, we gave an approach to show that how communication channel attack can be removed by using invisible watermarking. For insertion of the watermark we used the semi – pixel difference method. The watermarked template is sent to matching module which is present abroad. The watermark is extracted at the matching module site and at that site secret template is checked for its validity by using the extracted watermark. In future, we will try to improve the security of biometrics further by using data hiding techniques and cryptography.

References

[1] yadav,R.K.; chawla,G and Saini, R, "Semi Pixel Difference Method For Digital Image Watermarking With Minimum Degradation In Image Quality"

[2] Jain, A.K.; Uludag, U., "Hiding biometric data", Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 25, Issue: 11, Nov. 2003.

[3] F. Hartung and M. Kutter, "Multimedia Watermarking Techniques," Proc, IEEE, vol. 87, no. 7, pp. 1079-1107, July 1999.

[4] M. Kutter, F. Jordan, and F. Bossen, "Digital Signature of Color Images Using Amplitude Modulation," Proc. SPIE, vol. 3022, pp. 518-526, 1997.

[5] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT Domain System for Robust Image Watermarking," Signal Processing, vol. 66, no. 3, pp. 357-372, May 1998.

[6] Mikdam A. T. Alsalami and Marwan M. AL- Akaidi,” Digital Audio Watermarking”

[7] Kant,C;Nath,R;and chaudhary,S”Biometrics Security using Steganography” International Journal of Security, Volume (2) : Issue (1),2008

[8] Soutar C., “Biometric system security,” White Paper, Bioscrypt, http://www.bioscrypt.com..2004

[9] Uludag U., Pankanti S., Prabhakar S., and Jain A. K., “Biometric cryptosystems: issues and challenges,” Proceedings of the IEEE, vol. 92, no. 6, pp. 948–960, 2004.

[10] Jain, A.K., Bolle, R., and Pankanti S.,”Biometrics: Personal Identification in Networked Society”, Kluwer Academic Publishers, 1999.

[11] Jain A. K., Ross A., and Prabhakar S., “An introduction to biometric recognition,” IEEE Trans. Circuits Syst. Video Technology, Special Issue Image- and Video-Based Biomet., Volume 14, Issue 1, Jan. 2004, pp. 4–20.

[12] Maltoni D., Maio D., Jain A. K., and Prabhakar. S. Handbook of Fingerprint Recognition. Springer, New York, 2003.

[13] Jain, A.K., Uludag, U., “Hiding biometric data”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 25, Issue: 11, Nov. 2003.

[14] Jain, A.K., Uludag, U., “Hiding biometric data”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 25, Issue: 11, Nov. 2003.

[15] Ratha N. K., Connell J. H., and R. M. Bolle. An analysis of minutiae matching strength. In Proc. AVBPA 2001, Third International Conference on Audio- and Video-Based Biometric Person Authentication, pages223.

[16] Ratha, N.K., Connell J.H., and R.M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems”, IBM Systems Journal, vol. 40, no. 3, 2001.

[17] Waldmann, Ulrich, Dirk S., and Claudia E., “Protected transmission of biometric user authentication data for on card-matching,” Proceedings of the 2004 ACM symposium on Applied computing March 2004.

[18] Jain, Anil K. and Arun Ross, “Multibiometric systems,” Communications of the ACM,” January 2004, Volume 47,Number 1 (2004).

[19] Anil. K., Ross A., and Prabhakar S., “An introduction to biometric recognition,” IEEE Trans. Circuits Syst. Video Technology, Special Issue Image- and Video-Based Biomet., Volume 14, Issue 1, Jan. 2004, pp. 4–20.

[20] Ratha, N.K., Connell J.H., and R.M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems”, IBM Systems Journal, vol. 40, no. 3, 2001.

[21] Ambalakat, P. ”Security of Biometric Authentication Systems”, 21st Computer Science Seminar SA1-T1-1

PERSONALIZATION CONCEPT BASED USER PROFILE ON SEARCH ENGINE

D. Nagesh[1]
[1]PG Scholar
Department of IT, Aurora's Engineering College,
Bhongir, Nalgonda, Andhra Pradesh, India

M. V. Vijaya Saradhi[2]
[2]Professor& HOD
Department of IT, Aurora's Engineering College,
Bhongir, Nalgonda, Andhra Pradesh, Indi

Abstract: In the present days the search engines return the same results for the same query, regardless of the user's interest. So in order to avoid that we introduce new concept called personalization. The Personalized Search aims to customize search results according to each individual user for him to find the most relevant documents to him on the top by considering his idiosyncrasies. This could possibly satisfy them and help in finding relevant information easily and quickly. User profile is a component of any personalization applications. Most existing user profiling strategies are based on objects that users like, but not the objects that users dislike. In this paper, we focus on search engine personalization and develop concept-based user profiling methods that are based on both preferences. Users can be mined from the concept-based user profiles to perform mutual filtering. Browsers with same idea and can share their knowledge

Keywords: User Profile, Personalization, Clustering, Query logs, Ranking, Search Engine.

I.INTRODUCTION

There has been a tremendous growth in the amount of information on the web. Information retrieval systems are critical for overcoming this information overload and providing the information of interest to users of the systems. Users typically pose a short query consisting of a few keywords describing their information need. Information Retrieval systems perform a 'word to word' match of the query words with all documents in their document collection and return documents containing the words entered. Retrieval in a web scenario is much harder due to the large and dynamic content on the web.

Major web search engines usually cater to hundreds of millions of users and hundreds of millions queries every day. It is very unlikely that the millions of users are similar in interests and search for similar information. Also, it is probable that the query words entered by users exhibit polysemy (same word used in different senses like 'Java' can be used to mean Java the programming language or Java islands in Indonesia) and synonymy (different words can be used to convey similar information like OOP and Object Oriented Programming) due to ambiguous nature of natural language. Therefore, given different backgrounds of users, different interests of users and ambiguities in natural language, it is very likely that query words of two different users may appear exactly same even though information needs are different. However, current retrieval systems perform a 'word to word' match of the query words and work in a "one size fits all" fashion using the same search procedure for all the users. This makes the current retrieval systems far from optimal.

This inherent non-optimality is seen clearly in the following three cases: [1] When a query contains ambiguous terms: Different users may use exactly the same query (e.g., "Java") to search for different information (e.g., the Java island in Indonesia or the Java programming language), but existing IR systems return the same results for these users. Without considering the actual user, it is impossible to know which sense "Java" refers to in a query. [2] When a query contains partial information: A query can contain an acronym or a shorter usage of a longer phrase. Then there

might not be sufficient information required to infer information need of user. For example a query like "SBH" can mean "State Bank of Hyderabad" or "Syracuse Behavioral Healthcare" among others. Existing IR systems return mixture of results containing the exact word which might contain different expansions. Knowledge of interests and/or location of the user could be helpful in gathering more information required to understand the query. [3] When information need of the user changes: A users information needs may change over time. The same user may use "Java" sometimes to mean the Java island in Indonesia and some other times to mean the programming language. Without recognizing the search context, it would be again impossible to recognize the correct sense. Thus using user context information about user and query is necessary for improving the retrieval performance. Indeed, personalized search essentially boils down to capturing and exploiting related user context information of a query to improve search accuracy.

II.RELATED WORK

In the today's world because of advancement in the technology and increase in the computer literacy, computer and internet are becoming the most necessary part of human life. And people are using search engines to search necessary information. The search engine has the different type of users.

A. A brief history of web searching

Search engines as we know them today began to appear in 1994 when the number of HTTP resources increased. However, Internet search engines were in use before the emergence and growth of the Web. The first pre-Web search engine was Archie, which allowed keyword searches of a database of names of files available via FTP. The first robot and search engine of the Web was Wandex, which was developed by Matthew Gray in 1993. Since the appearance and exponential growth of the Web, hundreds of search engines with different features have appeared.

Primary search engines were designed based on traditional information retrieval methods. AltaVista, Lycos and Excite made huge centralized indices of Web pages. To answer a query, they simply retrieved results from their indexed databases and showed the cached pages based on keyword occurrence and proximity. While traditional indexing models have been successful in databases, it was revealed that these methods are not sufficient for a tremendously unstructured information resource such as the Web. The completeness of the index is not the only factor in the quality of search results. "Junk results" often wash out any results that a user is interested in. In order to increase the quality of search, Google made an innovative ranking system for the entire Web. PageRank used the citation graph of the Web and Google introduced link analysis in the search engine systems. Other efforts have been made to customize and specialize search tools.

Current retrieval systems (or search engines) return a long list of results obtained by 'word to word' match with query words. However, it has been observed that users typically view only top few (usually [10]) documents out of the long list of results returned by search engines. This requires retrieval systems to show the most relevant documents to a user on the top to improve user satisfaction with the search engine. However, without knowledge about the user context, this task is di cult to do because"relevance" of a document depends on the individual user and the individual query.

	Short term (dynamic)	Long term (static)
Explicit	immediately judged relevant document	hobbies, occupation interests
Implicit	immediately clicked document	query log

Table 1.1: Classification and examples of User Context

III.PERSONALIZATION SEARCH

A. Conceptual Based Search

Most concept-based methods automatically derive users' topical interests by exploring the contents of the users' browsed documents and search histories. Liu et al. [13] proposed a user profiling method based on users' search history and the Open Directory Project (ODP) [16]. The user profile is represented as a set of categories, and for each category, a set of keywords with weights. The categories stored in the user profiles serve as a context to disambiguate user queries. If a profile shows that a user is interested in certain categories, the search can be narrowed down by providing suggested results according to the user's preferred categories.

Gauch et al. [9] proposed a method to create user profiles from user browsed documents. User profiles are created using concepts from the top four levels of the concept hierarchy created by Magellan [14]. A classifier is employed to classify user browsed documents into concepts in the reference ontology. Xu et al. [20] proposed a scalable method which automatically builds user profiles based on users' personal documents (e.g. browsing histories and emails). The user profiles summarize users' interests into

hierarchical structures. The method assumes that terms exist frequently in user's browsed documents represent topics that the user is interested in. Frequent terms are extracted from users' browsed documents to build hierarchical user profiles representing users' topical interests.

Liu et al. and Gauch et al. both use reference ontology (e.g. ODP) to develop the hierarchical user profiles, while Xu et al. automatically extracts possible topics from users' browsed documents and organizes the topics into hierarchical structures. The major advantage of dynamically building a topic hierarchy is that new topics can be easily recognized and extracted from documents and added to the topic hierarchy, whereas reference ontology such as ODP is not always up to-date. Thus, all of the proposed users profiling strategies rely on a concept extraction method, which extracts concepts from web-snippets2 to create accurate and up-to-date user profiles.

Fig1.Concept Based Search

B. Document Based Search

Most document-based methods focus on analyzing users' clicking and browsing behaviors recorded in the user's clickthrough data. On web search engines, clickthrough data is an important implicit feedback mechanism from users. Clickthrough data for the query "apple", which contains a list of ranked search results presented to the user, with identification on the results that the user has clicked on.

Joachim's [10] proposed a method which employs preference mining and machine learning to model users' clicking and browsing behavior. Joachim's method assumes that a user would scan the search result list from top to bottom. If a user has skipped a document d_i at rank i before clicking on document d_j at rank j, it is assumed that he/she must have scan the document d_i and decided to skip it. Thus, it can be concluded that the user prefers document dj more than document d_i (i.e. $d_j < r' d_i$, where r' is the user's preference order of the documents in the search result list).

C. Query logs

Search Query logs consist of logs of searches made by users of search engines. They are usually collected at the search engine server. They typically consist of : user identity (ip address or anonymous id etc), search queries, corresponding clickthroughs made by the user and click information regarding it like the click time, no of clicks made etc. Some times the query logs are also captured on the client side i.e., on the user's computers. Clickthrough data/Query logs have been the most important source for capturing user context for user modeling. There has been some work in this connection some of which are described below.

In [6], Sugiyama et. al used web browsing history in past N days for personalized search. They partition the browsing history data into three categories according to the time stamp, i.e., persistent data (before today), today data (today but before the current session) and current session data. They found that the performance of using web browsing history is competitive with that using relevance feedback. Speretta et al[11] also used users search history to construct user profiles. Several other works have made use of past queries mined from the query logs to help the current searcher. (see [[7], [8], [9], [10]]).
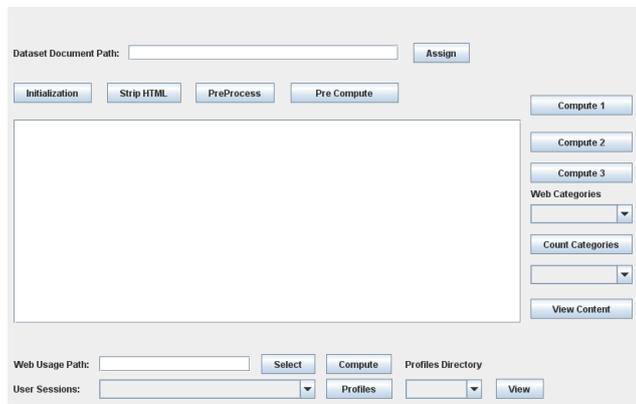
Fig2 .Concept Based Search Assigning the Databases

D. Query Clustering Algorithm

Concept-based user profiles are employed in the clustering process to achieve personalization effect.

First, a query-concept bipartite graph G is constructed by the clustering algorithm with one set of nodes corresponds to the set of users' queries, and the other corresponds to the sets of extracted concepts. Each individual query submitted by each user is treated as an individual node in the bipartite graph by labeling each query with a user identifier. Concepts with interestingness weights (defined in Equation 1) greater than zero in the user profile are linked to the query with the corresponding interestingness weight in G.

Second, a two-step personalized clustering algorithm is applied to the bipartite graph G, to obtain clusters of similar queries and similar concepts. The personalized clustering algorithm iteratively merges the most similar pair of query nodes, and then the most similar pair of concept nodes, and then merges the most similar pair of query nodes, and so on. The following cosine similarity function is employed to compute the similarity score sim(x, y) of a pair of query nodes or a pair of concept nodes. The advantages of the cosine similarity are that it can accommodate negative concept weights and produce normalized similarity values in the clustering process.

$$sim(x, y) = \frac{N_x \bullet N_y}{\|N_x\|\|N_y\|} \quad (1)$$

The algorithm is divided into two steps, initial clustering and community merging. In initial clustering, queries are grouped within the scope of each user. Community merging is then involved to group queries for the community. A more detailed example is provided in our previous work [11] to explain the purpose of the two steps in our personalized clustering algorithm

A common requirement of iterative clustering algorithms is to determine when the clustering process should stop to avoid over-merging of the clusters. Likewise, a critical issue in Algorithm 1 is to decide the termination points for initial clustering and community merging. When the termination point for initial clustering is reached, community merging kicks off; when the termination point for community merging is reached, the whole algorithm terminates.
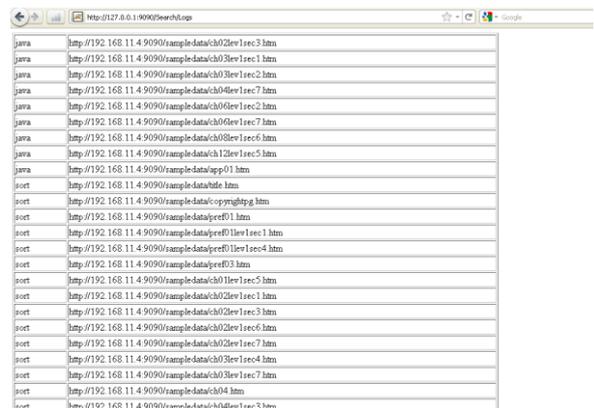
Fig 3. User Logs

Good timing to stop the two phases is important to the algorithm, since if initial clustering is stopped too early (i.e., not all clusters are well formed), community merging merges all the identical queries from different users, and thus generates a single big cluster without much personalization effect. However, if initial clustering is stopped too late, the clusters are already overly merged before community merging begins. The low precision rate thus resulted would undermine the quality of the whole clustering process
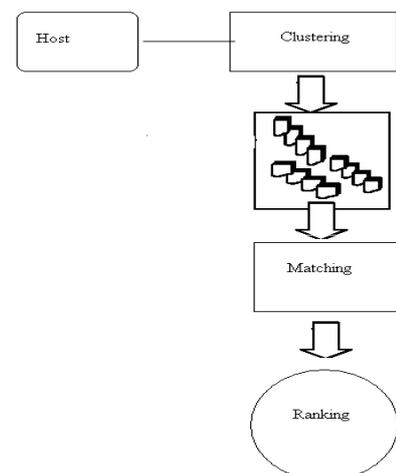
Fig 4. Document Clustering.

IV.USER PROFILING

This section proposes two user profiling strategies which are both concept-based and utilize users' positive and negative preferences.

They are
A. PJoachims−C
B. PClick+Joachims−C

A. Joachims-C Method (P$_{Joachims}$−C)

Joachims' original method was based on users' document preferences. If a user has skipped a document d_i at rank i before clicking on document d_j at rank j, he/she must have scanned the document d_i and decided to skip it. Thus, we can conclude that the user prefers document dj more than document d_i (i.e., $d_j < r'\ d_i$, where r' is the user's preference order of the documents in the search result list).

It is extended Joachims' method, which is a document-based method, to a concept based method (Joachims-C). Instead of obtaining the document preferences $d_j < r'\ d_i$, Joachims-C assumes that the user prefers the concepts $C(d_j)$ associated with document d_j to the concepts $C(d_i)$ associated with document d_i, and produces the corresponding concept preferences. From this it can be concluded that the concepts $C(d5)$ is more relevant to the user than the concepts in the other three unclicked documents (i.e., $C(d2)$, $C(d3)$ and $C(d4)$). The concept preference pair's extracted using Joachims-C method is shown in Table 4.2.

Concept Preference Pairs for d_1	Concept Preference Pairs for d_5	Concept Preference Pairs for d_8
Empty Set	apple store $<_{r'}$ product macintosh $<_{r'}$ product	macintosh $<_{r'}$ product catalog $<_{r'}$ product
	apple store $<_{r'}$ mac os macintosh $<_{r'}$ mac os	macintosh $<_{r'}$ mac os catalog $<_{r'}$ mac os
	macintosh $<_{r'}$ apple store apple store $<_{r'}$ iPod macintosh $<_{r'}$ iPod	macintosh $<_{r'}$ apple store catalog $<_{r'}$ apple store macintosh $<_{r'}$ iPod catalog $<_{r'}$ iPod
		macintosh $<_{r'}$ fruit catalog $<_{r'}$ fruit macintosh $<_{r'}$ apple hill catalog $<_{r'}$ apple hill
		macintosh $<_{r'}$ fruit catalog $<_{r'}$ fruit

Table 4.1 Concept Preference Pairs Obtained Using Joachims-C Methods

After the concept preference pairs are identified using Proposition , a ranking SVM algorithm [10] is employed to learn the user's preferences, which is represented as a weighted concept vector. Given a set of concept preference pairs T, ranking SVM aims at finding a linear ranking function f(q, c) to rank the extracted concepts so that as many concept preference pairs in T as possible are satisfied. f(q, c) is defined as the inner product of a weight vector −→w and a feature vector of query-concept mapping φ(q, c), which describes how well a concept c matches the user's interest for a query q.

B. Click+Joachims-C Method (PClick+Joachims−C)

In [11], it is observed that PClick is good in capturing user's positive preferences. In this paper, it is integrated the click-based method, which captures only positive preferences, with the Joachims-C method, with which negative preferences can be obtained. It is found that Joachims-C is good in predicting users' negative preferences. Since both the user profiles PClick and PJoachims−C are represented as weighted concept vectors, the two vectors can be combined using the following formula:

$$W(c+j)cs = w(c)cs + w(j)cs \quad if\ w(j)cs < 0$$
$$W(c+j)cs = w(c)cs\ other\ wise \text{ ----------------1}$$
$$W(c+j)ci = w(c)ci + w(j)ci \quad if\ w(j)c < 0 \text{------2}$$

where $w(C + J)ci \in PClick+Joachims−C$, $w(C)ci \in PClick$, and $w(J)ci \in PJoachims−C$. If a concept ci has a negative weight in $PJoachims−C$ (i.e., $w(J)ci < 0$), the negative weight will be added to $w(C)ci$ in PClick (i.e., $w(J)ci + w(C)ci$) forming the weighted concept vector for the hybrid profile $PClick+Joachims−C$.
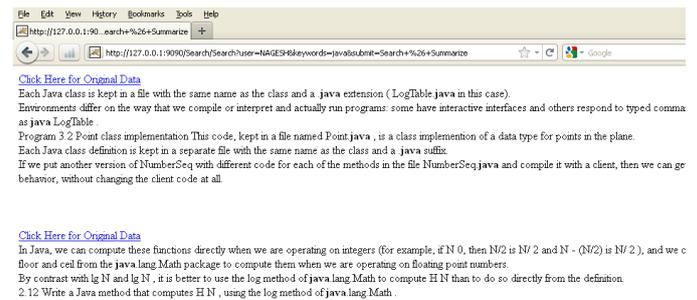
Fig4. User Personalized Based Search Results
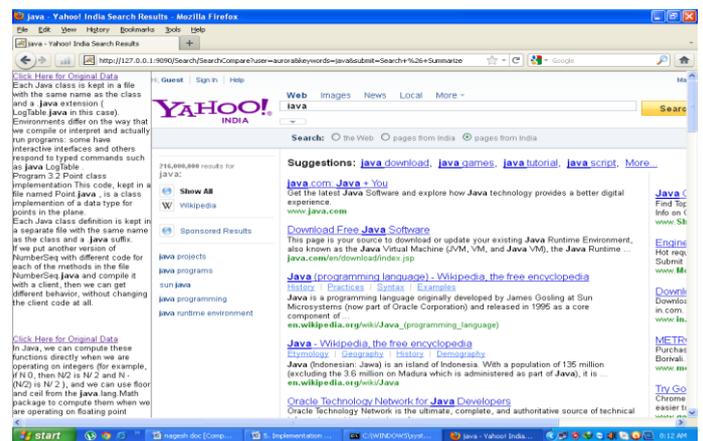
Fig5. User Personalized Comparison Search Result with yahoo

V.CONCLUSION

In this paper it has been presented a user profile strategy to improve a search engine's performance by identifying the information needs for individual users. For clustering of documents both content based clustering and session based clustering techniques is used. To automate the identification of groups of similar pages, the approach has been implemented in a Java prototype. This paper

proposes an effective method for organizing and visualizing web search results. Finally, the concept-based user profiles can be integrated into the ranking algorithms of a search engine so that search results can be ranked according to individual users' interests. To automate the identification of groups of similar pages, the approach has been implemented in a Java prototype. This paper proposes an effective method for organizing and visualizing web search results.

REFERENCES

[1]. P. Ingwersen and N. Belkin. Information retrieval in context irix. SIGIR Forum,38(2), 2004.

[2.].J. J. Rocchio. Relevance feedback in information retrieval, the smart retrieval system. Experiments in Automatic Document Processing, pages 313–323, 1971.

[3].Thorsten Joachims. Optimizing search engines using clickthrough data. In KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 133–142, New York, NY, USA, 2002. ACM Press.

[4].D Kelly and J Teevan. Implicit feedback for inferring user preference: A bibliog-raphy. In SIGIR Forum, volume 32, 2003.

[5].M. Claypool, M. Waseda P. Le, and D. Brown. Implicit interest indicators. In Proceedings of Intelligent User Interfaces 2001, pages 33–40, 2001.

[6].K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on userprofile constructed without any e_ort from users. In Proceedings of WWW 2004,pages 675 – 684, 2004.

[7].Vijay V. Raghavan and Hayri Sever. On the reuse of past optimal queries. Pro-ceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 344–350, 1995.

[8].Jian-Yun Ji-Rong Wen and Hong-Jiang Zhang. Query clustering using user logs.ACM Transactions on Information Systems (TOIS), 20(1):59–81, 2002.

[9].Larry Fitzpatrick and Mei Dent. Automatic feedback using past queries: Social searching? In In Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 306–313. ACM Press, 1997.

[10].Natalie S. Glance. Community search assistant. In In Proceedings of the International Conference on Intelligent User Interfaces, pages 91–96. ACM Press, 2001.

[11].Micro Speretta and Susan Gauch. Personalizing search based on user search histories. In Thirteenth International Conference on Information and Knowledge Management (CIKM 2004), 2004.

[12].J. Teevan, S. T. Dumais, and E. Horvitz. Personalizing search via automated analysis of interests and activites. In Proceedings of SIGIR 2005, 2005.

[13].Filip Radlinski and Thorsten Joachims. Evaluating the robustness of learning from implicit feedback. In ICML Workshop on Learning In Web Search, 2005.

[14].Boris Chidlovskii, Nathalie Glance, and Antonietta Grasso. Collaborative re-ranking of search results. In Proc. AAAI-2000 Workshop on AI for Web Search., 2000.

[15].Henxi Lin., Gui-Rong Xue., Hua-Jun Zeng., and Yong Yu. Using probabilistic latent semantic analysis for personalized web search. In Proceedings of APWEB'05, 2005.

[16].Armin Hust. Query expansion methods for collaborative information retrieval.Inform., Forsch. Entwickl., 19(4):224–238, 2005.

Authors

D.NAGESH is currently pursuing M.Tech (WT) at Aurora's Engineering College, Bhongir, Andhra Pradesh, India.

Dr. M.V.Vijaya Saradhi received his Ph.D degree from Faculty of Engineering, Osmania University (OU), Hyderabad, Andhra Pradesh, India**.** He is Currently Working as Professor in the Department of Information Technology (IT) at Aurora's Engineering College, Bhongiri, Andhra Pradesh, India. His main research interests are Software Metrics, Distributed Systems, Object-Oriented Modeling, Data Mining, Design Patterns, Object- Oriented Design Measurements and Empirical Software Engineering. He is a life member of various Professional bodies like MIETE, MCSI, MIE, MISTE.

A PRAGMATIC TECHNIQUE FOR DETECTION AND REMOVAL OF CRACKS IN DIGITIZED PAINTINGS

Pranob K Charles[1] ,K Balaji[1], V.V.S. Murthy[1],Rajendra Prasad.K[2], Sripath Roy K[2], Purnachand S[2], Suresh A[2]

[1]Assoc.Prof's, Dept of ECE, K.L University, [2]Asst. Prof's, Dept of ECE, K.L University

ABSTRACT

A Pragmatic technique for the detection and removal of cracks on digitized paintings is presented in this paper. The cracks are detected by thresholding the output of the morphological top-hat and bottom-hat transforms. And the outputs of the crack detection stage are compared to know which transform gives better results. Finally, crack filling using order statistics filters is done. These techniques have shown to perform very well on digitized paintings suffering from cracks.

KEY WORDS- *Detection of cracks, order statistics filters, top-hat transform, bottom-hat transform, virtual restoration of paintings.*

I. INTRODUCTION

We generally see paintings, especially old ones; suffer from breaks in the substrate, the paint, or the varnish. These patterns are usually called cracks or craquelure and can be caused by aging, drying, and mechanical factors. Age cracks can result from non uniform contraction in the canvas or wood-panel support of the painting, which stresses the layers of the painting. Drying cracks are usually caused by the evaporation of volatile paint components and the consequent shrinkage of the paint. Finally mechanical cracks result from painting deformations due to external causes, e.g., vibrations and impacts. The appearance of cracks on paintings deteriorates the image quality. However, one can use digital image processing techniques to detect and eliminate the cracks on digitized paintings. Such a "virtual" restoration can provide clues to art historians, museum curators and the general public on how the painting would look like in its initial state, i.e., without the cracks. Furthermore, it can be used as a non destructive tool for the planning of the actual restoration. The user should manually select a point on each crack to be restored. A method for the detection of cracks using transformation techniques is discussed in crack detection phase. Other research areas that are closely related to crack removal include image in painting which deals with the reconstruction of missing or damaged image areas by filling in information from the neighbouring areas, and disocclusion, i.e., recovery of object parts that are hidden behind other objects within an image. Methods developed in these areas assume that the regions where information has to be filled in are known. Different approaches for interpolating information in structured and textured image areas have been developed. The former are usually based on partial differential equations (PDEs) and on the calculus of variations whereas the latter rely on texture synthesis principles. A technique that decomposes the image to textured and structured areas and uses appropriate interpolation techniques depending on the area where the missing information lies has also been proposed. The results obtained by these techniques are very good. Different techniques for the restoration of cracks on digitized paintings, which adapts and integrates a number of image processing and analysis tools is proposed in this paper. The technique consists of the following stages:

• *crack detection;*
• *crack filling (interpolation).*

User interaction is rather unavoidable since the large variations observed in the typology of cracks would lead any fully automatic algorithm to failure. However, all processing steps can be executed in real time, and, thus, the user can instantly observe the effect of parameter tuning on the image under study and select in an intuitive way the values that achieve the optimal visual result. The results obtained after restoration of deteriorated images was very positive. This paper is organized as follows. Section II describes the crack-detection procedure. Methods for filling the cracks with image content from neighbouring pixels are proposed in Section III. Conclusions and discussion follow.

II. DETECTION OF CRACKS

Cracks usually have low luminance and, thus, can be considered as local intensity minima with rather elongated structural characteristics. Therefore, a crack detector can be applied on the luminance component of an image and should be able to identify such minima. A crack-detection procedure based on the top-hat and bottom-hat transform is proposed in this paper. The major part is the structuring element and it is a shape, used to probe or interact with a given image, with the purpose of drawing conclusions on how this shape fits or misses the shapes in the image. It is typically used in morphological operations, such as dilation, erosion, opening, and closing, as well as the hit-or-miss transform. According to Georges Matheron, knowledge about an object (e.g., an image) depends on the manner in which we probe (observe) it. In particular, the choice of a certain s.e. for a particular morphological operation influences the information one can obtain. There are two main characteristics that are directly related to s.e.s:

- **Shape**. For example, the s.e. can be a ``ball'' or a line; convex or a ring, etc. By choosing a particular s.e., one sets a way of differentiating some objects (or parts of objects) from others, according to their shape or spatial orientation.

- **Size**. For example, one s.e. can be a 3×3 square or a 21×21 square. Setting the size of the structuring element is similar to setting the observation scale, and setting the criterion to differentiate image objects or features according to size.

In structuring element there are two major concepts they are **HIT** and **FIT:**

FIT: All on pixels in the structuring element cover on pixels in the image

HIT: Any on pixel in the structuring element covers an on pixel in the image

The top hat transform performs opening operation i.e erosion followed by dilation:

$$A \circ B = (A \ominus B) \oplus B \qquad (1)$$

Here A is the image and B is the structuring element.
It is nothing but union operation performed on the structuring element and the image

$$A \circ B = \bigcup_{B_x \subseteq A} B_x \qquad (2)$$

The two major concepts are erosion and dilation - Erosion of image *f* by structuring element *s* is given by

$$f \ominus s$$

The structuring element s is positioned with its origin at *(x, y)* and the new pixel value is determined using the rule:

$$g(x, y) = \begin{cases} 1 \text{ if } s \text{ fits } f \\ 0 \text{ otherwise} \end{cases}$$

Fig 1: Example of erosion.

Dilation of image *f* by structuring element *s* is given by

$$f \oplus s$$

The structuring element s is positioned with its origin at *(x, y)* and the new pixel value is determined using the rule:

$$g(x, y) = \begin{cases} 1 \text{ if } s \text{ hits } f \\ 0 \text{ otherwise} \end{cases}$$

Fig 2: Example of dilation.

And bottom-hat transform is an operation that extracts small elements and details from given images. Then in this closing of images takes place. Closing operation is nothing but dilation and then erosion.

$$A \bullet B = (A \oplus B) \ominus B \qquad (3)$$

Here A is the image and B is the structuring element.
It is nothing but intersection operation performed on the structuring element and the image

$$A \bullet B = (A^c \circ B^s)^c \qquad (4)$$

Where X^c denotes the complement of X relative to E.

$$X^c = \{x \in E | x \notin X\} \qquad (5)$$

The top-hat transform generates a greyscale output image t(k,l) where pixels with a large grey value are potential crack or crack-like elements. Therefore, a thresholding operation on t(k,l) is required to separate cracks from the rest of the image. The threshold can be chosen by a trial and error procedure,i.e., by inspecting its effect on the resulting crack map. The low computational complexity of the thresholding operation enables the user to view the crack-detection results in real time while changing the threshold value, e.g., by moving a slider. This fact makes interactive threshold selection very effective and intuitive. Alternatively, threshold selection can be done by inspecting the histogram t(k,l) of for a lobe close to the maximum intensity value (which will most probably correspond to crack or crack-like pixels), and assigning it a value that separates this lobe from the rest of the intensities. The result of the thresholding is a binary image b(k,l) marking the possible crack locations. Instead of this global thersholding technique, more complex thresholding schemes, which use a spatially varying threshold, can be used. Obviously, as the threshold value increases the number of image pixels that are identified as cracks decreases. Thus, certain cracks, especially in dark image areas where the local minimum condition may not be satisfied, can remain undetected. In principle, it is more preferable to select the threshold so that some cracks remain undetected than to choose a threshold that would result in the detection of all cracks but will also falsely identify as cracks, and subsequently modify, other image structures. The thresholded (binary) output of the top-hat transform on the luminance component of an image containing cracks.And thus we can complete the crack detection stage.

III. CRACK-FILLING METHODS
After identifying cracks the final task is to restore the image using local image information (i.e., information from neighboring pixels) to fill (interpolate) the cracks. Two classes of techniques, utilizing order statistics

filtering and anisotropic diffusion are proposed for this purpose. Both are implemented on each RGB channel independently and affect only those pixels which belong to cracks. Therefore, provided that the identified crack pixels are indeed crack pixels, the filling procedure does not affect the "useful" content of the image. Image in painting techniques like the ones cited in Section I can also be used for crack filling. The performance of the crack filling methods presented below was judged by visual inspection of the results. The results have been verified and the images are being restored.

3.1 Crack Filling Based on Order Statistics Filters

An effective way to interpolate the cracks is to apply median or other order statistics filters in their neighborhood. All filters are selectively applied on the cracks, i.e., the center of the filter window traverses only the crack pixels. If the filter window is sufficiently large, the crack pixels within the window will be outliers and will be rejected. Thus, the crack pixel will be assigned the value of one of the neighboring non crack pixels.

The following filters can be used for this purpose.

• Median filter

$$y_i = \mathrm{med}(x_{i-\nu}, \ldots, x_i, \ldots, x_{i+\nu}) \tag{6}$$

• Recursive median filter

$$y_i = \mathrm{med}(y_{i-\nu}, \ldots, y_{i-1}, x_i, \ldots, x_{i+\nu}) \tag{7}$$

Where the $y_{i-\nu, \ldots\ldots, y_{i-1}}$ are the already computed median output samples. For both the recursive median and the median filter, the filter window (considering only rectangular windows) should be approximately 50% wider than the widest (thickest) crack appearing on the image.This is necessary to guarantee that the filter output is selected to be the value of a noncrack pixel. Smaller windows will result in cracks that will not be sufficiently filled whereas windows that are much wider than the cracks will create large homogeneous areas, thus distorting fine image details.

• Weighted median filter

$$y_i = \mathrm{med}(w_{-\nu} \Diamond x_{i-\nu}, \ldots, w_{\nu} \Diamond x_{i+\nu}) \tag{8}$$

Where $w \Diamond x$ denotes duplication of times. For this filter, smaller filter windows (e.g., windows that are approximately 30% wider than the widest crack appearing on the image) can be used since the probability that a color value corresponding to a crack is selected as the filter output (a fact that would result in the crack pixel under investigation not being filled effectively by the filter) can be limited by using small weights for the pixels centrally located within the window (which are usually part of the crack) and bigger ones for the other pixels.

The filter coefficients are chosen as follows:

$$\alpha_{rs} = \begin{cases} 0, & \text{if } \mathrm{med}\{x_{ij}\} - x_{i+r, j+s} \geq q \\ 1, & \text{otherwise} \end{cases} . \tag{10}$$

The amount of trimming depends on the positive parameter. Data of small value deviating strongly from the local median (which correspond usually to cracks) are trimmed out. Windows used along with this variant of the MTM filter can also be smaller than those used for the median and recursive median filters since a portion of the crack pixels is expected to be rejected by the trimming procedure.

•And after finding the threshold output which contain only binary values for the pixels i.e either 1 or 0.All the pixels with value 1 indicate non cracked part of the image and all the pixel with values 0 indicate cracked pixels where the data is missing. Now we collect the matrix columns and rows of the pixel with value 0 by using the condition:

$$\text{if } b(i,j) = 0 \tag{11}$$

Now after obtaining the row and column values of the pixels where the data is missing we now use the mean value of the neighborhood pixels as shown below:

Table 1 : Neighbouring Pixels of cracked pixel

B1	B2	B3
B4	B5	B6
B7	B8	B9

Here in the above table let B5 is the cracked pixel or the pixel with value 0 and now we calculate the mean of the neighbor pixels using:

Mean=(B1+ B2+ B3+ B4+ B5+ B6+ B7+ B8+ B9)/N

$$\tag{12}$$

N=no of pixels .

Now after we get the mean value and we substitute the mean value in place of the missing data likewise we perform the same operation on all the missing pixel values and we get the missing data and we can reconstruct the image.

The median operator can be used instead of the arithmetic mean in (9). And after finding the assigning the pixel For this variant of the MTM filter, even smaller filter windows can be used, since crack pixels do not contribute to the filter output. Thus, it suffices that the window is 1 pixel wider than the widest crack. The result of the application of the second variation of the modified trimmed mean filter on the painting depicted in Fig. 3 (filter size 5x5). Another image restored by the same crack-filling approach can be seen in Fig. 4 (filter size 3x3). Extensive experimentation proved that this filter gives the best results among all filters presented above according to the evaluations. The superiority of this filter can be attributed to the fact that only non crack pixels contribute to its output. Thus the cracks can be interpolated.

IV RESULTS

Fig3: Original image

*Fig6:*original painting
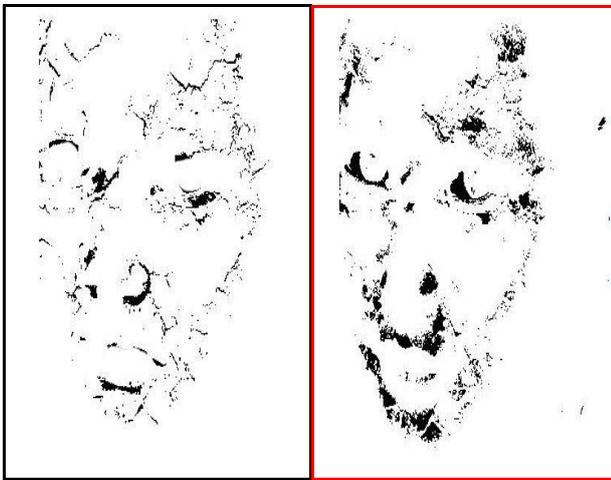
Fig 4 a) Threshold output of bottom hat transform
b) Threshold output of top hat transform
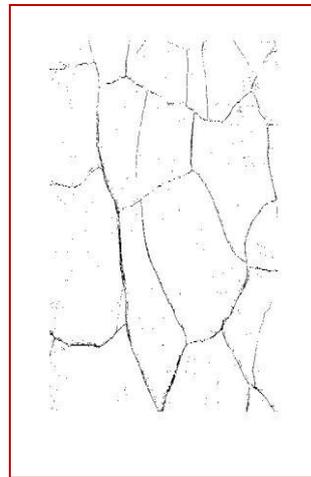
Fig7: output after applying transform and thresholding.

Fig5: Cracks Interpolated and reconstructed image.

Fig8: Reconstructed Image.

V. CONCLUSIONS AND DISCUSSION

In this paper, we have presented advanced techniques for detection and filling in digitized paintings. Cracks are detected by using top-hat & bottom hat transforms and compared the results and observed that bottom hat transform yields better results. Crack interpolation is performed by appropriately modified order statistics filters. The methodology has been applied for the virtual restoration of images and was found very effective. However, there are certain aspects of the proposed methodology that can be further improved. For example, the crack-detection stage is not very efficient in detecting cracks located on very dark image areas, since in these areas the intensity of crack pixels is very close to the intensity of the surrounding region. A possible solution to this shortcoming would be to apply the crack-detection algorithm locally on this area and select a low threshold value. Another situation where the system (more particularly, the crack filling stage) does not perform as efficiently as expected is in the case of cracks that cross the border between regions of different colour. In such situations, it might be the case that part of the crack in one area is filled with colour from the other area, resulting in small spurs of colour in the border between the two regions. However, this phenomenon is rather seldom and, furthermore, the extent of these erroneously filled areas is very small (2–3 pixels maximum). A possible solution would be to perform edge detection or segmentation on the image and confine the filling of cracks that cross edges or region borders to pixels from the corresponding region. Use of image in painting techniques could also improve results in that aspect. Another improvement of the crack filling stage could aim at using properly adapted versions of nonlinear multichannel filters (e.g., variants of the vector median filter) instead of processing each colour channel independently. These improvements will be the topic of future work on this subject. This can be implemented for the pavement crack detection which has very importance now a days.

REFERENCES

[1] M. Barni, F. Bartolini, and V. Cappellini, "Image processing for virtual restoration of artworks," IEEE Multimedia, vol. 7, no. 2, pp. 34–37, Jun 2000.

[2] F. Abas and K. Martinez, "Craquelure analysis for content-based retrieval," in Proc. 14th Int. Conf. Digital Signal Processing, vol. 1, 2002,pp. 111–114.

[3] I. Giakoumis, N. Nikolaidis, and L Pitas,Digital lmage Processing Techniques for the Detection and Removal of Cracks in Digitized Paintings, IEEE Transactions on Image ProcessingV, ol . l5, No.I , pp. l78-l 88. Jan. 2006.

[4] A. Kokaram, R. Morris, W. Fitzgerald, and P. Rayner, "Detection of missing data in image sequences," IEEE Trans. Image Process., vol. 4, no. 11, pp. 1496–1508, Nov. 1995.

[5] "Interpolation of missing data in image sequences," IEEE Trans.Image Process., vol. 4, no. 11, pp. 1509–1519, Nov. 1995.

[6] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in Proc. SIGGRAPH, 2000, pp. 417–424.

[7] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera"Filling-in by joint interpolation of vector fields and gray levels," IEEE Trans. Image Process., vol. 10, no. 8, pp. 1200–1211, Aug. 2001.

[8] S. Masnou and J. M. Morel, "Level lines based disocclusion," in Proc.IEEE Int. Conf. Image Process., vol. III, 1998, pp. 259–263.

[9] L. Joyeux, O. Buisson, B. Besserer, and S. Boukir, "Detection and removal of line scratches in motion picture films," in Proc. IEEE Int. Conf.Computer Vision and Pattern Recognition, 1999, pp. 548–553.

[10] D. King, The Commissar Vanishes, Henry Holt and Company, 1997.

[11] S. Masnou and J.M. Morel, Level-lines based disocclusion, 5th IEEE International Conference on Image Processing, Chicago, IL, Oct 4–7, 1998.

[12] M. Nitzberg, D. Mumford, and T. Shiota, Filtering, Segmentation, and Depth, Springer-Verlag, Berlin, 1993.

[13] S. Walden, The Ravished Image, St. Martin's Press, New York, 1985.

[14] G. Emile-Male, The Restorer's Handbook of Easel Painting, Van Nostrand Reinhold, New York, 1976.

[15] G. R. Arce, "Multistage order statistic filters for image sequence processing," IEEE Trans. Signal Processing, vol. 39, pp. 1146-1 161, May 1991.

[16] Bilge Alp, Petri Haavisto, Tiina Jarske, Kai Oistamo, and Yrjo Neuvo, "Median-based algorithms for image sequence processing," SPIE Visual Commun. Image Processing, 1990, pp. 122-133.

[17] A. C. Kokaram and P. J. W. Rayner, "A system for the removal of impulsive noise in image sequences," in SPIE Visual Commun. Image Processing, Nov. 1992, pp. 322-331.

[18] A. C. Kokaram, "Motion picture restoration," Ph.D. thesis, Cambridge Univ., UK, May 1993.

[19] S. Efstratiadis and A. Katsagellos, "A model-based, pel-recursive motion estimation algorithm," in Proc. IEEE ICASSP, 1990, pp. 1973-1976.

[20] J. Konrad and E. Dubois, "Bayesian estimation of motion vector fields," IEEE Trans. Patt. Anal. Machine Intell,, vol. 14, no. 9, Sept. 1992.

DETECTION OF E-BANKING PHISHING WEBSITES

E.Konda Reddy[1], Dr. Rajamani[2] and Dr. M. V. Vijaya Saradhi[3]

[1]PG Scholar, [2]Dean Informatics & Professor and [3]Head Of the Department and Professor
Department of IT, Aurora Engineering College, Bhongir, Nalgonda, Andhra Pradesh, India.

Abstract- Phishing is a new type of network attack where the attacker creates a replica of an existing web page to fool users in to submitting personal, financial, or password data to what they think is their service provider's website. The concept is an end-host based anti-phishing algorithm, called the Link Guard, by utilizing the generic characteristics of the hyperlinks in phishing attacks. The link Guard algorithm is the concept for finding the phishing emails sent by the phisher to grasp the information of the end user. Link Guard is based on the careful analysis of the characteristics of phishing hyperlinks. Each end user is implemented with Link Guard algorithm. After doing so the end user recognizes the phishing emails and can avoid responding to such mails. Since Link Guard is characteristics based it can detect and prevent not only known phishing attacks but also unknown ones. The project uses the Java technologies and Oracle.

Keywords- Phishing, Fuzzy Logic, Data Mining, Classification, association, e-banking risk assessment

I.INTRODUCTION

Phishing is a new word produced from 'fishing', it refers to the act that the attacker allure users to visit a faked Web site by sending them faked e-mails (or instant messages), and stealthily get victim's personal information such as user name, password, and national security ID, etc.

This information then can be used for future target advertisements or even identity theft attacks (e.g., transfer money from victims' bank account). The frequently used attack method is to send e-mails to potential victims, which seemed to be sent by banks, online organizations, or ISPs. In these e-mails, they will make up some causes, e.g. the password of your credit card had been mis-entered for many times, or they are providing upgrading services, to allure you visit their Web site to conform or modify your account number and password through the hyperlink provided in the e-mail.

If you input the account number and password, the attackers then successfully collect the information at the server side, and is able to perform their next step actions with that information (e.g., withdraw money out from your account).Phishing itself is not a new concept, but it's increasingly used by phishers to steal user information and perform business crime in recent years. Within one to two years, the number of phishing attacks increased dramatically. Our analysis identifies that the phishing hyperlinks share one or more characteristics as listed below:

1) The visual link and the actual link are not the same;
2) The attackers often use dotted decimal IP address instead of DNS name;
3) Special tricks are used to encode the hyperlinks maliciously;

4) The attackers often use fake DNS names that are similar (but not identical) with the target Web site.

We then propose an end-host based anti-phishing algorithm which we call Link Guard, based on the characteristics of the phishing hyperlink. Since Link Guard is character-based, it can detect and prevent not only known phishing attacks but also unknown ones. We have implemented Link Guard in Windows XP, and our experiments indicate that Link Guard is light-weighted in that it consumes very little memory and CPU circles, and most importantly, it is very effective in detecting phishing attacks with minimal false negatives.

The paper is organized as follows: Section 2 presents the literature review and related work. Section 3 presents the existing anti phishing approaches. Section 4 introduces the system design and implementation of Link Guard approach. and then conclusions and future work are given in Section 5.

II. LITERATURE REVIEW AND RELATED WORK

A. Literature Review

Phishing website is a recent problem, nevertheless due to its huge impact on the financial and on-line retailing sectors and since preventing such attacks is an important step towards defending against e-banking phishing website attacks, there are several promising approaches to this problem and a comprehensive collection of related works. In this section, we briefly survey existing anti-phishing solutions and list of the related works. One approach is to stop phishing at the email level [3], since most current phishing attacks use broadcast email (spam) to

lure victims to a phishing website [19]. Another approach is to use security toolbars. The phishing filter in IE7 [18] is a toolbar approach with more features such as blocking the user's activity with a detected phishing site. Other approach is to visually differentiate the phishing sites from the spoofed legitimate sites. Dynamic Security Skins [5] proposes to use a randomly generated visual hash to customize the browser window or web form elements to indicate the successfully authenticated sites. A fourth approach is two- factor authentication, which ensures that the user not only knows a secret but also presents a security token [6]. However, this approach is a server-side solution. Phishing can still happen at sites that do not support two-factor authentication. Sensitive information that is not related to a specific site, *e.g.*, credit card information and SSN, cannot be protected by this approach either [20].

However, an automatic anti-phishing method is seldom reported. The typical technologies of anti-phishing from the User Interface aspect are done by [5] and [20]. They proposed methods that need Web page creators to follow certain rules to create Web pages, either by adding dynamic skin to Web pages or adding sensitive information location attributes to HTML code. However, it is difficult to convince all Web page creators to follow the rules [7].

B. Main Characteristics of e-banking phishing websites.

Evolving with the anti-phishing techniques, various phishing techniques and more complicated and hard-to-detect methods are used by phishers. The most straightforward way for a phisher to defraud people is to make the phishing Web pages similar to their targets. Actually, there are many characteristics and factors that can distinguish the original legitimate website from the forged e-banking phishing website like Spelling errors, Long URL address and Abnormal DNS record. The full list is shown in table I which will be used later on our analysis and methodology study.

III. EXISTING SYSTEM
We briefly review the approaches for anti-phishing.

1) Detect and block the phishing Web sites in time:
If we can detect the phishing Web sites in time, we then can block the sites and prevent phishing attacks. It's relatively easy to (manually) determine whether a site is a phishing site or not, but it's difficult to find those phishing sites out in time. Here we list two methods for phishing site detection.

Table I. COMPONENTS AND LAYERS OF E-BANKING PHISHING WEBSITE CRITERIA.

Criteria	N	Component	Layer No.
URL & Domain Identity (Weight = 0.3)	1	Using the IP Address	Layer One
	2	Abnormal Request URL	
	3	Abnormal URL of Anchor	
	4	Abnormal DNS record	Sub weight = 0.3
	5	Abnormal URL	
Security & Encryption (Weight = 0.2)	1	Using SSL certificate	Layer Two
	2	Certification authority	
	3	Abnormal Cookie	
	4	Distinguished Names Certificate(DN)	
Source Code & Java script (Weight = 0.2)	1	Redirect pages	Sub weight = 0.4
	2	Straddling attack	
	3	Pharming Attack	
	4	Using onMouseOver to hide the Link	
	5	Server Form Handler (SFH)	
Page Style & Contents (Weight =0.1)	1	Spelling errors	Layer Three
	2	Copying website	
	3	Using forms with "*Submit*" button	
	4	Using Pop-Ups windows	
	5	Disabling Right-Click	
Web Address Bar (Weight = 0.1)	1	Long URL address	Sub weight = 0.3
	2	Replacing similar characters for URL	
	3	Adding a prefix or suffix	
	4	Using the @ Symbol to Confuse	
	5	Using Hexadecimal Character Codes	
Social Human Factor (Weight = 0.1)	1	Much emphasis on security and response	
	2	Public generic salutation	
	3	Buying Time to Access Accounts	

A) The Web master of a legal Web site periodically scans the root DNS for suspicious sites (e.g. www. 1 cbc.com.cn vs. www.icbc.com.cn).
B) Since the phisher must duplicate the content of the target site, he must use tools to (automatically) download the Web pages from the target site.

It is therefore possible to detect this kind of download at the Web server and trace back to the phisher. Both approaches have shortcomings. For DNS scanning, it increases the overhead of the DNS systems and may cause problem for normal DNS queries, and furthermore, many phishing attacks simply do not require a DNS name. For phishing download detection, clever phishers may easily write tools which can mimic the behavior of human beings to defeat the detection.

2) Enhance the security of the web sites:
The business Websites such as the Web sites of banks can take new methods to guarantee the security of users' personal information. One method to enhance the security is to use hardware devices. For example, the Barclays bank provides a hand-held card reader to the users. Before shopping in the net, users need to insert their credit card into the card reader, and input

their (personal identification number) PIN code, then the card reader will produce a onetime security password, users can perform transactions only after the right password is input. Another method is to use the biometrics characteristic (e.g. voice, fingerprint, iris, etc.) for user authentication. For example, PayPal had tried to replace the single password verification by voice recognition to enhance the security of the Web site.

With these methods, the phishers cannot accomplish their tasks even after they have gotten part of the victims' information. However, all these techniques need additional hardware to realize the authentication between the users and the Web sites hence will increase the cost and bring certain inconvenience. Therefore, it still needs time for these techniques to be widely adopted.

3) Block the phishing e-mails by various spam filters: Phishers generally use e-mails as 'bait' to allure potential victims. SMTP (Simple Mail Transfer Protocol) is the protocol to deliver e-mails in the Internet. It is a very simple protocol which lacks necessary authentication mechanisms. Information related to sender, such as the name and email address of the sender, route of the message, etc., can be counterfeited in SMTP. Thus, the attackers can send out large amounts of spoofed e-mails which are seemed from legitimate organizations. The phishers hide their identities when sending the spoofed e-mails, therefore, if anti-spam systems can determine whether an e-mail is sent by the announced sender (Am I Whom I Say I Am?), the phishing attacks will be decreased dramatically.

From this point, the techniques that preventing senders from counterfeiting their Send ID (e.g. SIDF of Microsoft) can defeat phishing attacks efficiently. SIDF is a combination of Microsoft's Caller ID for E-mail and the SPF (Sender Policy Framework) developed by Meng Weng Wong. Both Caller ID and SPF check e-mail sender's domain name to verify if the e-mail is sent from a server that is authorized to send e-mails of that domain and from that to determine whether that e-mail use spoofed e-mail address. If it's faked, the Internet service provider can then determine that e-mail is a spam e-mail. The spoofed e-mails used by phishers are one type of spam e-mails. From this point of view, the spam filters can also be used to filter those phishing e-mails. For example, blacklist, white list, keyword filters, Bayesian filters with self learning abilities, and E-Mail Stamp, etc., can all be used at the e-mail server or client systems. Most of these anti-spam techniques perform filtering at the receiving side by scanning the contents and the address of the received e-mails. And they all have pros and cons as discussed below. Blacklist and whitelist cannot work if the names of the

spammers are not known in advance. Keyword filter and Bayesian filters can detect spam based on content, hence can detect unknown spasm. But they can also result in false positives and false negatives. Furthermore, spam filters are designed for general spam e-mails and may not very suitable for filtering phishing e-mails since they generally do not consider the specific characteristics of phishing attacks.

4) Install online anti-phishing software in user's computers: Despite all the above efforts, it is still possible for the users to visit the spoofed Web sites. As a last defense, users can install anti-phishing tools in their computers. The antiphishing tools in use today can be divided into two categories: blacklist/white list based and rule-based.

Category I: When a user visits a Web site, the antiphishing tool searches the address of that site in a blacklist stored in the database. If the visited site is on the list, the anti-phishing tool then warns the users. Tools in this category include Scam Blocker from the EarthLink Company, Phish Guard, and Net craft, etc. Though the developers of these tools all announced that they can update the blacklist in time, they cannot prevent the attacks from the newly emerged (unknown) phishing sites.

Category II: this category of tools uses certain rules in their software, and checks the security of a Web site according to these rules. Examples of this type of tools include Spoof Guard developed by Stanford, Trust Watch of the Geo Trust, etc. Spoof Guard checks the domain name, URL (includes the port number) of Web site, it also checks whether the browser is directed to the current URL via the links in the contents of e-mails. If it finds that the domain name of the visited Web site is similar to a well-known domain name, or if they are not using the standard port, Spoof Guard will warn the users. In Trust Watch, the security of a Web site is determined by whether it has been reviewed by an independent trusted third party organization. Both Spoof Guard and Trust Watch provide a toolbar in the browsers to notify their users whether the Web site is verified and trusted.

It is easy to observe that all the above defense methods are useful and complementary to each other, but none of them are perfect at the current stage.

IV. PROPOSED SYSTEM AND ITS IMPLEMENTATION

In this section we explain the basic algorithm of Link Guard Approach which can detect the phishing content, based on the characteristics of the phishing hyperlink.

LINKGUARD

A.Classification of the hyperlinks in the phishing e-mails

In order to (illegally) collect useful information from potential victims, phishers generally tries to convince the users to click the hyperlink embedded in the phishing e-mail. A hyperlink has a structure as follows.

 Anchor text <\a>

where 'URI' (universal resource identifiers) provides the necessary information needed for the user to access the networked resource and 'Anchor text' is the text that will be displayed in user's Web browser. Examples of URIs are
http://www.google.com,
https://www.icbc.com.cn/login.html,
ftp://61.112.1.90:2345, etc. 'Anchor text' in general is used to display information related to the URI to help the user to better understand the resources provided by the hyperlink. In the following hyperlink, the URI links to the phishing archives provided by the APWG group, and its anchor text "Phishing Archive" informs the user what's the hyperlink is about.

<a   href"http://www.antiphishing.org/phishing archive.html">
Phishing Archive

Note that the content of the URI will not be displayed in user's Web browser. Phishers therefore can utilize this fact to play trick in their 'bait' e-mails. In the rest of the paper, we call the URI in the hyperlink the actual link and the anchor text the visual link. After analyzing the 203 (there are altogether 210 phishing e-mails, with 7 of them with incomplete information or with malware attachment and do not have hyperlinks) phishing email archives from Sep. 21st 2003 to July 4th 2005 provided by APWG [6]. We classified the hyperlinks used in the phishing e-mail into the following categories:

1) The hyperlink provides DNS domain names in the anchor text, but the destination DNS name in the visible link doesn't match that in the actual link. For instance, the following hyperlink:
<a   href   =
"http://www.profusenet.net/checksession.php">
https://secure.regionset.com/EBanking/logon/
appears to be linked to secure.regionset.com, which is the portal of a bank, but it actually is linked to a phishing site www.profusenet.net.

2) Dotted decimal IP address is used directly in the URI or the anchor text instead of DNS name. See below for an example.

 SIGN IN

3) The hyperlink is counterfeited maliciously by using certain encoding schemes. There are two cases: a) The link is formed by encoding alphabets into their corresponding ASCII codes. See below for such a hyperlink.
 www.citibank.com
while this link is seemed pointed www.citibank.com, it actually points to http://4.34.195.41:34/l/index.htm.

b) Special characters (e.g. @ in the visible link) are used to fool the user to believe that the e-mail is from a trusted sender. For instance, the following link seems is linked to amazon, but it actually is linked to IP address 69.10.142.34.
http://www.amazon.com:fvthsgbljhfcs83infoupdate @69.10.142.34.

4) The hyperlink does not provide destination information in its anchor text and uses DNS names in its URI. The DNS name in the URI usually is similar with a famous company or organization. For instance, the following link seems to be sent from paypal, but it actually is not. Since paypal-cgi is actually registered by the phisher to let the users believe that it has something to do with paypal

 Click here to confirm your account

5) The attackers utilize the vulnerabilities of the target Web site to redirect users to their phishing sites or to launch CSS (cross site scripting) attacks. For example, the following link
<a href="http://usa.visa.com/track/dyredir.jsp?rDirl=http://200.251.251.10/.verified/"> Click here <a>
Once clicked, will redirect the user to the phishing site 200.251.251.10 due to a vulnerability of usa.visa.com. Table 1 summarizes the number of hyperlinks and their percentages for all the categories. It can be observed that most of the phishing e-mails use faked DNS names (category 1,44.33%) or dotted decimal IP addresses (category 2, 41.87%).

Encoding tricks are also frequently used (category 3a and 3b, 17.24%). And phishing attackers often try to fool users by setting up DNS names that are very similar with the real ecommerce sites or by not providing destination information in the anchor text (category 4).

Phishing attacks that utilize the vulnerability of Web sites (category 5) are of small number (2%) and we leave this type of attacks for future study.

Note that a phishing hyperlink can belong to several categories at the same time. For instance, an attacker may use tricks from both categories 1 and 3 at the same time to increase his success chance. Hence the sum of percentages is larger than 1.

Category	Number of links	Percentage
1	90	44.33%
2	85	41.87%
3.a	19	9.36%
3.b	16	7.88%
4	6	7.33%
5	4	2%

TABLE 2
THE CATEGORIES OF HYPERLINKS IN PHISHING E-MAILS.

Once the characteristics of the phishing hyperlinks and understood, we are able to design anti-phishing algorithms that can detect known or unknown phishing attacks in real-time. We present our LinkGuard algorithm in the next subsection.

B. The LinkGuard algorithm
LinkGuard works by analyzing the differences between the visual link and the actual link. It also calculates the similarities of a URI with a known trusted site. The algorithm
is illustrated in Fig. 1. The following terminologies are used in the algorithm.
v_link: visual link;
a_link: actual_link;
v_dns: visual DNS name;
a_dns: actual DNS name;
sender_dns: sender's DNS name.
int LinkGuard(v_link, a_link} {
1 v_dns = GetDNSName(v_link);
2 a_dns = GetDNSName(a_link);
3 if ((v_dns and a_dns are not
4 empty) and (v_dns != a_dns))
5 return PHISHING;
6 if (a_dns is dotted decimal)
7 return POSSIBLE_PHISHING;
8 if(a_link or v_link is encoded)
9 {
10 v_link2 = decode (v_link);
11 a_link2 = decode (a_link);
12 return LinkGuard(v_link2, a_link2);
13 }
14 /* analyze the domain name for
15 possible phishing */
16 if(v_dns is NULL)
17 return AnalyzeDNS(a_link);
}
Fig. 1. Description of the LinkGuard algorithm.

The LinkGuard algorithm works as follows. In its main routine *LinkGuard*, it first extracts the DNS names from the actual and the visual links (lines 1 and 2). It then compares the actual and visual DNS names, if these names are not the same, then it is phishing of category 1 (lines 3-5). If dotted decimal IP address is directly used in actual dns , it is then
a possible phishing attack of category 2 (lines 6 and 7). We will delay the discussion of how to handle possible phishing attacks later. If the actual link or the visual link is encoded

int AnalyzeDNS (actual_link) {
/* Analyze the actual DNS name according
to the blacklist and whitelist*/
18 if (actual_dns in blacklist)
19 return PHISHING;
20 if (actual_dns in whitelist)
21 return NOTPHISHING;
22 return PatternMatching(actual_link);
}
int PatternMatching(actual_link){
23 if (sender_dns and actual_dns are different)
24 return POSSIBLE_PHISHING;
25 for (each item prev_dns in seed_set)
26 {
27 bv = Similarity(prev_dns, actual_link);
28 if (bv == true)
29 return POSSIBLE_PHISHING;
30 }
31 return NO_PHISHING;
}
float Similarity (str, actual_link) {
32 if (str is part of actual_link)
33 return true;
34 int maxlen = the maximum string
35 lengths of str and actual_dns;
36 int minchange = the minimum number of
37 changes needed to transform str
38 to actual_dns (or vice verse);
39 if (thresh<(maxlen-minchange)/maxlen<1)
40 return true
41 return false;
}
Fig. 2. The subroutines used in the LinkGuard algorithm.

(categories 3 and 4), we first decode the links, then recursively call LinkGuard to return a result (lines 8-13). When there is no destination information (DNS name or dotted IP address) in the visual link (category 5), LinkGuard calls AnalyzeDNS to
analyze the actual dns (lines 16 and 17). LinkGuard therefore handles all the 5 categories of phishing attacks.
AnalyzeDNS and the related subroutines are depicted in Fig.2. In AnalyzeDNS, if the actual dns name is

contained in the blacklist, then we are sure that it is a phishing attack (lines 18 and 19). Similarly, if the actual dns is contained in the

whitelist, it is therefore not a phishing attack (lines 20 and 21). If the actual dns is not contained in either whitelist or blacklist, PatternMatching is then invoked (line 22). PatternMatching is designed to handle unknown attacks (blacklist/whitelist is useless in this case). For category 5 of the phishing attacks, all the information we have is the actual link from the hyperlink (since the visual link does not contain

DNS or IP address of the destination site), which provide very little information for further analysis. In order to resolve this problem, we try two methods: First, we extract the sender email address from the e-mail. Since phishers generally try to fool users by using (spoofed) legal DNS names in the sender e-mail address, we expect that the DNS name in the sender address will be different from that in the actual link. Second, we proactively collect DNS names that are manually input by the user when she surfs the Internet and store the names into a seed set, and since these names are input by the user by hand, we assume that these names are trustworthy. PatternMatching then checks if the actual DNS name of a hyperlink is different from the DNS name in the sender's address (lines 23 and 24), and if it is quite similar (but not identical) with one or more names in the seed set by invoking the Similarity (lines 25-30) procedure. Similarity checks the maximum likelihood of actual dns and the DNS names in seed set. As depicted in Fig. 2, the similarity index between two strings are determined by calculating the minimal number of changes (including insertion, deletion, or revision of a character in the string) needed to transform a string to the other string. If the number of changes is 0, then the two strings are identical; if the number of changes is small, then they are of high similarity; otherwise, they are of low similarity. For example, the similarity index of 'microsoft' and 'micr0s0ft' is 7/9 (since we need change the 2 '0's in micr0s0ft to 'o'. Similarly, the similarity index of 'paypal' and 'paypal-cgi' is 6/10 (since we need to remove the last 4 chars from paypal-cgi), and the similarity index of '95559'

and '955559' is 5/6 (since we need to insert a '5' to change '95559' to '955559').

If the two DNS names are similar but not identical, then it is a possible phishing attack. For instance, PatternMatching can easily detect the difference between www.icbc.com.cn (which is a good e-commerce Web site) and www.1cbc.com.cn (which is a phishing site), which has similarity index 75%. Note that PatternMatching may treat www.1cbc.com.cn as

a normal site if the user had never visit www.1cbc.com.cn before. This false negative, however, is unlikely to cause any severe privacy or

financial lose to the user, since she actually does not have anything to lose regarding the Web site www.icbc.com.cn (since she never visits that Web site before)!

C. False positives and false negatives handling

Since LinkGuard is a rule-based heuristic algorithm, it may cause false positives (i.e., treat non-phishing site as phishing site) and false negatives (i.e., treat phishing site as nonphishing site). In what follows, we show that LinkGuard may result in false positives but is very unlikely to cause harmful false negatives.

For phishing attacks of category 1, we are sure that there is no false positives or false negatives, since the DNS names of the visual and actual links are not the same. It is also easy to observe that LinkGuard handles categories 3 and 4 correctly since the encoded links are first decoded before further analysis. For category 2, LinkGuard may result in false positives, since using dotted decimal IP addresses instead of domain names may be desirable in some special circumstances (e.g., when the DNS names are still not registered). For category 5, LinkGuard may also result in false positives. For example, we know that both 'www.iee.org' and 'www.ieee.org' are legal Web sites. But these two DNS names have a similarity index of 3/4, hence is very likely to trigger a false positive.

When it is a possible false positive, LinkGuard will return a POSSIBLE PHISHING. In our implementation (which will be described in the next section), we leverage the user to judge if it is a phishing attack by prompting a dialogue box with detailed information of the hyperlink. The rationale behind this choice is that users generally may have more knowledge of a link than a computer in certain circumstances (e.g., the user may know that the dotted decimal IP address is the address of his friend's computer and that www.iee.org is a respected site for electrical engineers).

For category 5, LinkGuard may also result in false negatives. False negatives are more harmful than false positives, since attackers in this case will succeed in leading the victim to the phishing sites. For instance, when the sender's e-mail address and the DNS name in the actual link are the same and the DNS name in the actual link has a very low similarity index with the target site, LinkGuard will return NO PHISHING. For instance, PatternMatching will treat the below link as NO PHISHING.

```
<a href="http://fdic-secure.com/
application.htm"> Click here </a>
```

with "securehq@fdic-secure.com" as the sender address. We note that this kind of false negatives is very unlikely to result in information leakage, since the

end user is very unlikely to have information the attack interested (since the DNS name in this link is not similar with any legal Web sites).

V.   IMPLEMENTATION   AND VERIFICATION OF LINKGUARD

We have implemented the LinkGuard algorithm in Windows XP. It includes two parts: a whook.dll dynamic library and a LinkGuard executive. The structure of the implementation is depicted in Fig. 3.

This Link Guard algorithm is the concept for finding the phishing e-mails Sent by the phishers to grasp the information's of the end user. Link Guard is based the careful analysis of the characteristics of phishing hyperlinks. Link Guard has a verified very low false negative rate for unknown phishing attacks . This Link Guard algorithm is the concept for finding the phishing e-mails Sent by the phishers to grasp the information of the end user **,** So each end user will be implemented with the Link Guard algorithm , After implementing the Link Guard algorithm now the end user may able to find the phishing attacks, and can avoid responding phishing e-mails .
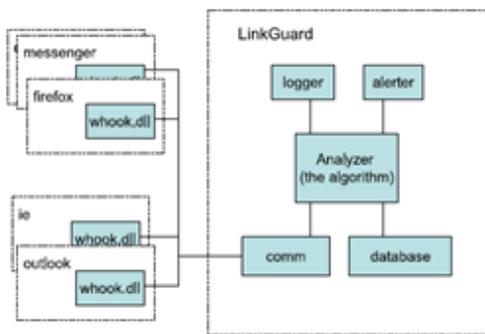

Fig. 3. The structure of the LinkGuard implementation, which consists of a whook.dll and a LinkGuard executive.

Since Link Guard is character-based, it can detect and prevent not only known phishing attacks but also unknown ones.
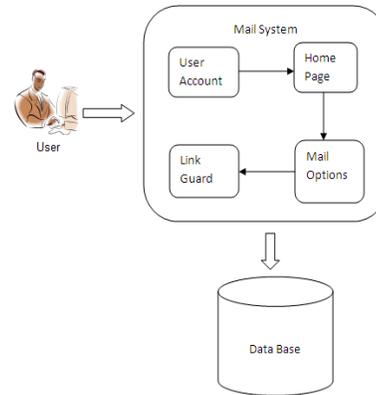
Figure 4: System Architecture of Anti phishing system along with Link Guard Approach

MODULES OF PROPOSED SYSTEM:

- Creation of a mail system and database   operations
- Composes, send and receive a mail
- Implementation of the Link Guard algorithm

The mail system module deals with the user interface for the home page, sign-in, sign-up and forgot your password pages. This module enables a new user to Sing-Up. It also enables an existing user to Sign-In. The user may use the Forget password link if he did forget his password. The password is retrieved on the basis of security question and answer given by the user. Database operation manages the users. Every time a new user signs in his details are written in to the database. Every time an existing user logs on his details are checked up for with the database.

The second module enables the user to compose and send a mail. It also allows the user to read a received mail. Once a mail is sent the date and the subject of the mail gets displayed. The received mail can be checked if it is phishing or not, the implementation of which is given in the next module. The compose mail option contains an option for spoof id. The spoof id allows the mail of the composer to be delivered with a different from address. This is being incorporated to demonstrate the Link Guard algorithm.

The module contains the implementation of the Link Guard algorithm. It is possible for the user to add domain names and categorize them as either white list or black list under settings. Whenever a mail is detected as phishing the domain name in that mail automatically gets added as black list. The Link Guard algorithm checks if the domain names fall under any of the 5 categories of hyperlinks for phishing emails. It also refers to the database of black and white list entries and sets the status of the mail as either **Phishing** or **Non-Phishing.**

Once the mail is categorized as Phishing the user can take care that he does not open the link or submit any personal, critical information on to the website.

Communicator: This collects the information of the input process, and sends these related information's to the Analyzer.

Database: Store the white list, blacklist, and the user input URLs.

Analyzer: It is the key component of Link Guard, which implements the Link Guard algorithm; it uses data provided by Communicator and Database, and sends the results to the Alert and Logger modules.

Alerter:When receiving warning messages from Analyzer, it shows the related information to alert the users and send back the reactions of the user back to the Analyzer.

Logger: Archive the history information, such as user events, alert information, for future use.

VI. CONCLUSION AND FUTURE WORK

Phishing has becoming a serious network security problem, causing finical lose of billions of dollars to both consumers and e-commerce companies. And perhaps more fundamentally, phishing has made e-commerce distrusted and less attractive to normal consumers. In this paper, we have studied the characteristics of the hyperlinks that were embedded in phishing e-mails.

We then designed an anti-phishing algorithm, Link-Guard, based on the derived characteristics. Since Link-Guard is characteristic based, it can not only detect known attacks, but also is effective to the unknown ones. We have implemented Link Guard for Windows XP. Our experiment showed that Link Guard is light-weighted and can detect up to 96% unknown phishing attacks in real-time. We believe that Link Guard is not only useful for detecting phishing attacks, but also can shield users from malicious or unsolicited links in Web pages and Instant messages.

As we have implemented this approach by considering the URL and Domain Identity Criteria, there are the different criteria needs to work in future.

REFERENCES

[1] Androutsopoulos, J. Koutsias, K.V Chandrinos, and C.D. Spyropoulos.An Experimental Comparison of Naive Bayesian and Keyword-Based Anti-Spam Filtering with Encrypted Personal E-mail Message.In Proc. SIGIR 2000, 2000.

[2] The Anti-phishing working group. http://www.antiphishing.org/.Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh, and John C.Mitchell. Client-side defense against web-based identity theft. In Proc.NDSS 2004, 2004.

[3] B. Adida, S. Hohenberger and R. Rivest, ―Lightweight Encryption for Email,‖ USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI), 2005.

[4] Cynthia Dwork, Andrew Goldberg, and Moni Naor. On Memory-Bound.Functions for Fighting Spam. In Proc. Crypto 2003, 2003.

[5] R. Dhamija and J.D. Tygar, ―The Battle against Phishing: Dynamic Security Skins,‖ Proc. Symp. Usable Privacy and Security, 2005.

[6] FDIC., ―Putting an End to Account-Hijacking Identity Theft,‖ http://www.fdic.gov/consumers/consumer/idtheftst udy/identity_theft.pdf, 2004.

[7] A. Y. Fu, L. Wenyin and X. Deng, ― Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover‗s Distance (EMD) ,‖ IEEE transactions on dependable and secure computing, vol. 3, no. 4, 2006.

[8] EarthLink. ScamBlocker. http://www.earthlink.net/software/free/toolbar/ .

[9] David Geer. Security Technologies Go Phishing. IEEE Computer, 38 (6):18-21, 2005.

[10] John Leyden. Trusted search software labels fraud site as safe'. http://www.theregister.co.uk/2005/09/27/untrusted-search/.

[11] Microsoft. Sender ID Framework. http://www.microsoft.com/

[12] mscorp/safety/technologies/senderid/default.mspx.

[13] Net craft. Net craft toolbar. http://toolbar.netcraft.com/.

[14] PhishGuard.com. Protect Against Internet Phishing Scams http://www.phishguard.com/.

[15] Jonathan B. Postel. Simple Mail Transfer Protocol. RFC821:http:llwww.ietf. org/rfc/rfcO82 1 .txt.

[16] Georgina Stanley. Internet Security - Gone phishing.http://www.cyota.com/news.asp?id=1 14.

[17] Meng Weng Wong. Sender ID SPF. http://www.openspf.org/whitepaper.pdf.

[18] T. Sharif, Phishing Filter in IE7, http://blogs.msdn.com/ie/archive/2005/09/09/4632 04.aspx, September 9, 2006.

[19] M. Wu, R. C. Miller and S. L. Garfinkel , ―*Do Security Toolbars Actually Prevent Phishing Attacks?,"* CHI April 2006.

[20] M. Wu, R. C. Miller and G. Little, ―Web Wallet: Preventing Phishing Attacks by Revealing User Intentions,‖ MIT Computer Science and Artificial Intelligence Lab, 2006.

Author Profile

E Konda Reddy, Pursuing M.Tech in the department of Information Technology, in Aurora's Engineerng College, Bhongir, Nalgonda Dist, Andhra Pradesh, India.

Dr.A.Rajamani, Professor, Dean Computer Science & Information Technology, He is currently working with Aurora's Engineering college, Bhongir, Nalgonda Dist, Andhra Pradesh, India.

Dr. M.V.Vijaya Saradhi received his Ph.D degree from Faculty of Engineering, Osmania University (OU), Hyderabad, Andhra Pradesh, India**. He** is Currently Working as Professor in the Department of Information Technology (IT) at Aurora's Engineering College, Bhongiri, Andhra Pradesh, India. His main research interests are Software Metrics, Distributed Systems, Object-Oriented Modeling, Data Mining, Design Patterns, Object- Oriented Design Measurements and Empirical Software Engineering. He is a life member of various Professional bodies like MIETE, MCSI, MIE, MISTE.

An Agent based Intrusion Detection System for Wireless Sensor Networks Using Multilevel Classification

K.Kulothungan[1], S.Ganapathy[2], P.Yogesh[3] and A.Kannan[4]

[1,2,3,4] Department of Information Science & Technology,
College of Engineering, Anna University, Chennai-25, Tamil nadu, India.

ABSTRACT

With the rapid growth of internet communication and availability of techniques to intrude the network, network security has become indispensable. In this paper, we propose a multilevel classification technique for intrusion detection that uses intelligent agents and a combination of decision tree classifier and Enhanced Multiclass Support Vector Machine algorithm for the implementation of an effective intrusion detection system in order to provide security to Wireless Sensor Networks. The main advantage of this approach is that the system can be trained with unlabeled data and is capable of detecting previously "unseen" attacks using agents. Verification tests have been carried out by using the KDD cup'99 data set. From the experiments conducted in this work, it has been observed that significant improvement has been achieved in intrusion detection rate and also in the reduction of false alarm rate.

Keywords*: Intrusion Detection, EMSVM, Multilevel Decision Tree, Intelligent Agents*

I.   INTRODUCTION

Recently, security has become a vital concern in many application areas since computers have been networked together with a very large number of users and systems. The adoption of Wireless Sensor Networks (WSNs) has increased in recent years mainly due to their advantages in many applications. WSNs can be defined as a heterogeneous system that consists of nodes which are having tiny sensors and actuators. Sensors networks may consist of hundreds or thousands of low-power, low-cost nodes, fixed nodes deployed largely together to monitor and affect the environment. Intrusion detection and prevention techniques are necessary to provide security to WSNs because they are prone to various types of attacks from both insiders and outsiders. Denial of Service (DoS) attacks is an important attack that leads to more power consumption as well as the collapse of the entire network due to unnecessary flooding of packets.

An intrusion detection system can be used as a first line of defense in such a scenario in order to reduce possible intrusions and thereby reducing the risk of attacks. The current security mechanisms such as firewalls focus only external attacks. On the other hand, intrusion detection systems are capable of detecting both internal and external attacks. IDSs are classified, based on their functionality as misuse and anomaly intrusion detection system. A misuse intrusion detection system uses a set of well defined patterns for attack and they can be detected by matching these patterns against normal user behavior in order to detect intrusions effectively. Usually, misuse detection is harder than the anomaly intrusion detection since it is carried out by legitimate internal users who know the systems password and other credentials. In an anomaly intrusion detection system, the user behavior becomes different from the normal usage behavior. Therefore, it is necessary to provide an intelligent intrusion detection system which can find out both internal and external attacks.

However, the existing intrusion detection techniques, which are proposed by various researchers for misuse and anomaly detection [14,2], are generally not sufficient to provide the required security to WSNs because they have limited power and tiny structure.   The attacks are carefully designed by the attackers and hence the application of the existing intrusion detection techniques causes a high false positive rate. Moreover, the existing Intrusion Detection techniques are capable of detecting only known intrusions since they classify instances by the rules they have acquired based on training from past data. However, it is necessary to build intelligent IDS with effective learning abilities in order to secure the network from both internal and external attacks.

In this paper, an intelligent agent based IDS that uses a multilevel classifier and also a decision maker agent for intelligently detecting the intruders in WSNs has been proposed and implemented so that it can provide effective security to WSNs. This intelligent system uses a combination of enhanced decision tree classifier and an Enhanced Multiclass SVM algorithm for binary classification of the past data as well as the current data.

We have enhanced the Support Vector Machines (SVM) for classification since SVM are the classifiers which are more effective in binary classification [19] [20]. In this work, we have combined SVMs with decision trees in order to design multiclass SVMs, which are capable of classifying the four types of attacks namely probing, DOS, U2R and R2L and normal data more accurately. The main focus of this paper is to provide a combined approach to detect the DDoS attacks which improves the training time, testing time and accuracy of IDS using this approach.

The reminder of this paper is organized as follows: Section 2 provides a survey of related works in the area of

misuse and anomaly detection, Decision trees and SVM. Section 3, depicts the architecture of the system proposed in this paper. Section 4 discusses about the proposed enhanced decision tree algorithm [12] and the enhanced multiclass SVM algorithm with agents. Section 5 shows the results obtained from this work and compares them with the existing works. Section 6 gives the conclusions on this work and suggests some possible future enhancements.

II. LITERATURE SURVEY

There are many works in the literature that deal with classification techniques [3] [16] [19]. For example, an algorithm called Tree structured Multiclass SVM has been proposed by Snehal A.Mulay et. al [15] for classifying data effectively. Their paper proposed decision tree based algorithms to construct multiclass IDS which are used to improve the training time, testing time and accuracy of IDS. However, the detection rate is not sufficient in the current internet scenario.

Multiple level tree classifiers were proposed by various researchers in the past [7] [8] [18] in order to design effective IDSs. In such systems, the data are split into normal DOS, PROBE and others (a new class label U2R and R2L). In the second level, the algorithm split the "others" into its corresponding U2R and R2L, while the third level classifies the attacks into its individual specific attacks. However, it is necessary to classify the DOS attacks with a special attention to improve the network performance.

Zeng and Wu et al [21] introduced a new anomaly detection approach based on multi-attribute decisional framework. The classification of data pattern is performed using K-nearest neighbour's method and SVM model. Experiments performed by them with KDD Cup 99 dataset demonstrate that their proposed method achieves good detection accuracy.

Kim and Reddy (2008) et al [22] introduced an anomaly IDS, which monitors packet headers of network traffic. It operates in postmortem but in real-time. The frequent attacks on network infrastructure, using various forms of DoS attacks, have led to an increased need for developing techniques for analyzing network traffic.

Cherkasova et al (2009) [23] proposed a novel framework that provides a powerful solution for automated anomaly detection and analysis of changes in application behavior. The online regression-based transaction model proposed in their work accurately detects a change in the Computational Power consumption pattern of the application and alarms about either observed performance anomaly or possible application change. One of the limitations of their work is that it cannot distinguish which of the transactions is responsible for a changed CPU consumption of the application. To complement the regression-based approach and to identify the transactions that cause the model change, they used the application performance signature that provides a compact model of runtime behavior of the application.

Techniques for the design and evaluation of Intrusion Detection models for Wireless Networks using a supervised classification algorithm and to evaluate the performance of the Multilayer Perceptron (MLP), and Support Vector Machine (SVM) has been provided by Aikaterini Mitrokotsa et. al [1].

The results provided by them point out that SVM exhibits high accuracy.

A novel architecture of Support Vector Machine classifiers utilizing binary decision tree (SVM -BDT) for solving the multiclass problems has been provided by Gjorgji Madzarov et. al [5]. This architecture provides techniques for achieving better classification accuracy.

In this paper, we propose an intelligent agent based multilevel classifier for IDS that uses a combination of decision tree classifier, enhanced C4.5 algorithm and intelligent agents for effective detection of intrusions in WSNs. This system applies the Enhanced Multiclass SVM algorithm for improving the training time, testing time and accuracy of IDS to reduce the false alarm rate. Comparing with existing works, the work proposed in this work different in many ways. First, this system uses intelligent agents for effective classification of DoS attacks. Second, this system uses a hybrid classification scheme for detecting intrusion. Finally, this system uses an enhanced C4.5 algorithm for effective classification.

III. SYSTEM ARCHITECTURE

The multilevel hybrid IDS architecture proposed in this paper is presented schematically in figure 1. This system consists of three modules where the tree classifier agent uses enhanced C4.5 algorithm with agent decision for constructing decision tree which is used to find misuse detection. The classification module uses Agent Multiclass SVM for unsupervised anomaly detection. Finally, for refined classification of anomaly detection, the agent based tree classifier has been used in this work.
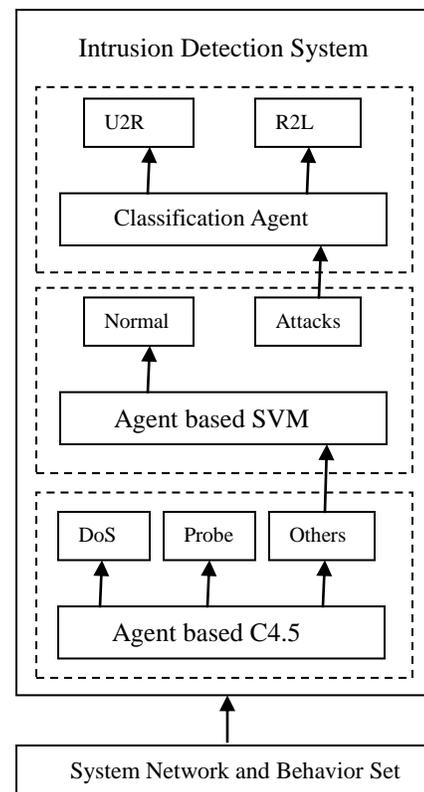

Figure 1 System Architecture

A. Classification Agent

This agent collects the KDD cup'99 data set and constructs a decision tree using the enhanced C4.5 algorithm where intelligent agents are used for making decision. The input data is mined and the specific attributes that have high information gain are only used in the construction of the decision tree where the KDD cup data is classified into DoS and the OTHERS categories. The DoS contains all specific Denial of service attacks like smurf, land, Neptune, back and teardrop. The others include the PROBE category as well. It contains the attacks including ipsweep, nmap, portsweep and Satan. All the other types of attacks and the normal connections are grouped into the OTHERS category.

The decision tree is first trained with the training data and decision tree is generated. The tree is pruned using agent to optimize the number of nodes in the decision tree. From the pruned decision tree, rules are formed. The rules are then applied to the test data and the input data is thus classified using the rules generated during the training phase.

- Stimulus/Response
  Stimulus: Collected training data
  Response: Decision tree with classified data

B. Behavior partition Module

In the Agent based Multiclass SVM algorithms, it is necessary to fix the number of classes are fixed in the beginning of classification. We use the agent based classification technology to determine the number of classes automatically. This Agent based Multiclass SVM algorithm is used to generate a classification whose outputs are normal and attack.

- Stimulus/Response

Stimulus: the part of the output from the decision tree classifier.
Response: two classes with are labeled normal and abnormal (attack).

IV. THE FRAMEWORK OF MULTILEVEL IDS

The main task of the Intrusion Detection System (IDS) is to discover the intrusions from the network packet data or system audit data. One of the major problems that the IDS might face is that the packet data or system audit data could be overwhelming. Moreover, some of the features of audit data may be redundant or contribute little to the detection process. Hence, agent based classification techniques has been used in this work to ease this task.

The network attacks fall into four main categories as discussed in [8].

• DoS (Denial of Service): Intrusions are designed to disrupt a host or network service, e.g. SYN flood;
• PROBE: Attacks include many programs which can automatically scan a network of computers to gather information or find known vulnerabilities as a possible precursor to more dangerous attacks.
• U2R (User to Root): Attacks correspond to a local user on a machine gaining privileges normally reserved for the UNIX root or super user.

• R2L (Remote to Local): Attacks correspond to an attacker who does not have an account on a victim machine, sends packets to that machine and gains local access, e.g. guessing password

In the next section, a brief introduction of the classification algorithms used in the hybrid IDS, i.e., the C4.5 algorithm for building decision trees and the Multiclass Support Vector Machine (SVM) are given.

Agent based Multiclass Support Vector Machine (EMSVM) Algorithm

In this section, we describe the intelligent Multiclass SVM algorithms, and illustrate how to apply this algorithm to generate anomaly type intrusion detection models. Figure 3 pictorially represents anomaly detection system discussed in this paper.

This agent based Multiclass Support Vector Machine (MSVM) algorithm is as follows: First, we first compute the distance between two classes of patterns and repeat it for each class of such patterns.
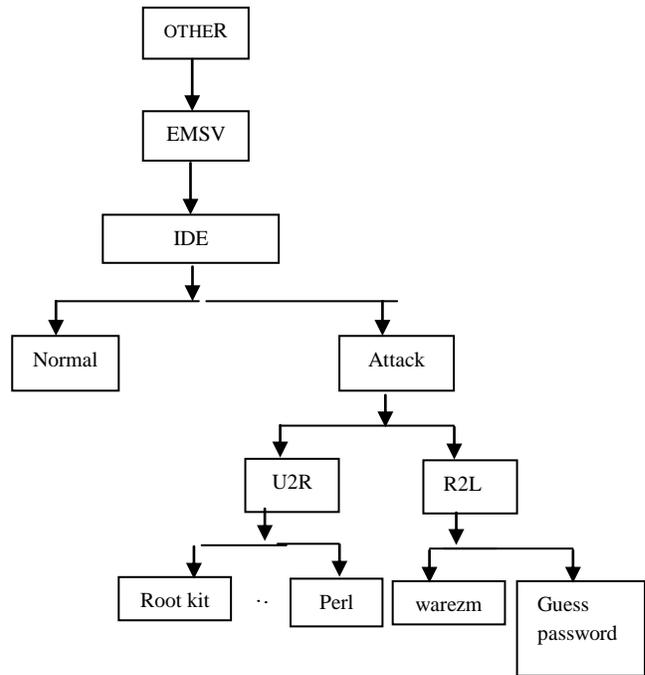
Figure 2. Behavior partition Module

where the distance between two classes is computed using the Minkowski Distance. According this method, the distance between two points

$$P = (x_1, x_2, \ldots, x_n) \text{ and } Q = (y_1, y_2, \ldots, y_n) \in R^n$$

is defined as

$$\left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$$

where p is the order.
We find the center point of every class by using the formula

$$C_i = \sum_{m=1}^{nt} X^i_m / n_i$$

After this calculation, five classes obtained earlier are converted into two classes. For example let A, B, C, D and E be five classes. If the distance between any two classes are less than that of the other classes then that pair is replaced by 1(Normal). Otherwise, it is replaced by -1 (Attacker). So, at end of the repeated process, we have only 1's and -1's combinations. Since -1 classes are removed, the remaining classes are used to construct the tree.

The steps of the algorithms are as follows:
Algorithm: Search (E, n).

Input: Data set E, the number of sampling.

Output: Initial center (m1, m2)

[1] Sampling E, get $S_1, S_2,…,S_n$

[2] For i=1 to n do

Mi =Count_m(Si);

[3] For i=1 to n do

M= Count_m(mi);

[4] m1=m, m2=max(Sim(m,m_i));

[5] Check it with agent threshold to make final descision.

Enhanced Multiclass Support Vector Machine algorithm

[1] Confirm two initial cluster centers by algorithm search m.

[2] Import a new class C.

[3] Compute the Minkowski distance between two classes.

[4] if (dAB > dAC) then
B is assigned as Normal
Else C is assigned as Attacker.

[5] Find the min & max of the distance.

[6] If (dAB < threshold limit of the distance) then create a new cluster and this is the center of the new cluster.
Else
B is assigned as an Attacker.

[7] Repeat the operation until reduced the difference between the classes.

[8] Validate this difference using agent.

V. EXPERIMENTATION AND RESULTS

A. Training and Test Data

The dataset used in the experiment was taken from the Third International Knowledge Discovery and Data Mining Tools Competition (KDD Cup 99). Each connection record is described by 41 attributes. The list of attributes consists of both continuous-type and discrete type variables, with statistical distributions varying drastically from each other, which makes the intrusion detection a very challenging task.

B. Experimental Results

Table 1 shows the comparison between C 4.5 and agent based C 4.5 with respect to DoS, Probe and Other types of attacks. From this table, it can be observed that the intrusion detection rate is improved in agent based C 4.5 when it is compared with the existing C 4.5 algorithm. This result was obtained by carrying out the experiments 20 times and then by

taking the average detection rate. The corresponding bar chart representation is shown in figure 3.

Table 1. Detection Rates (%) variation between C 4.5 and Agent based C4.5

Category	C 4.5	Agent based C 4.5
DoS	99.19	99.59
Probe	99.71	99.82
Others	66.67	68.06
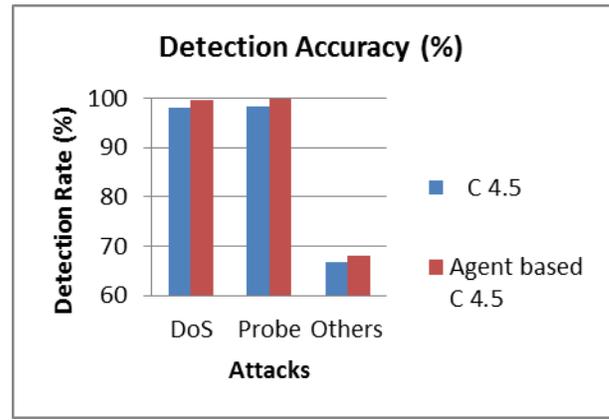
Figure 3 Comparison of Performance Analysis of C4.5 and Agent based C4.5

Table 2 shows the comparison of results obtained from Multiclass Support Vector machine(MSVM) and Agent based Multiclass Support Vector Machine algorithms. From this table, it is observed that the false positive rate has been reduced in the agent based MSVM when it is compared with MSVM. Moreover, the classification time is also reduced in the agent based MSVM due to the effective decisions made by the agent in classification.

Table 2. The Results Comparison between MSVM and AMSVM

Algorithm	TN	Accuracy (%)	R-error (%)	T-time (Sec)
MSVM	14756	83.5821	6.5147	846
Agent based MSVM	15332	92.4612	5.2136	223
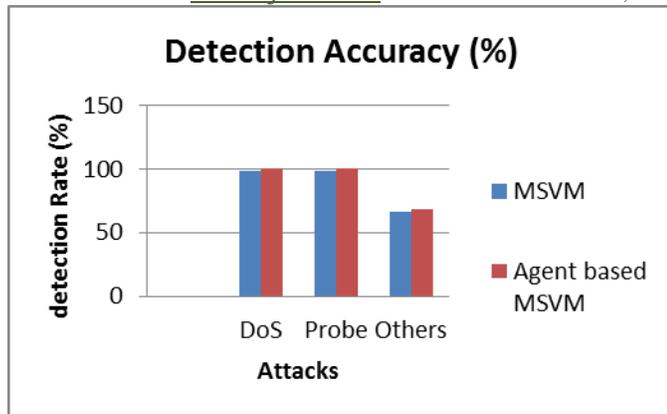
Figure 4. The Results Comparison between MSVM and Agent based MSVM

Figure 4 shows the comparison of intrusion detection rate between MSVM and Agent based MSVM algorithms. From this, it can be seen that the agent based MSVM performs well with respect to detection rate.

VI.   CONCLUSION AND FUTURE WORKS

In this paper, an intelligent intrusion detection system using an agent based multi-level classification model combining Decision trees and an agent based Multiclass Support Vector Machines has been proposed. From the experiments conducted in this work, it can be concluded that the agent based intrusion detection system improves the detection accuracy by 7% and 5% when it is compared with C 4.5 and SVM algorithms for DoS attacks. Moreover, it reduces the false positive rate by 1% when it is compared with existing system. Further works in this direction could be the use of effective preprocessing techniques for attribute selection in the IDS to enhance the performance.

REFERENCES

[1] Aikaterini Mitrokotsa, Manolis Tsagkaris and Christos Douligeris, "Intrusion *Detection in Mobile Ad hoc Networks using Classification Algorithms*", 2009.
[2] Denning D E, "*An Intrusion Detection Model*",. *IEEE Transactions on Software Engineering, Vol. 51, no. 8, pp. 12-26, Aug. 2003.*
[3] Denning D E, "*An Intrusion-Detection Model*", *IEEE Transaction on Software Engineering*, Vol. 13, No. 3, pp. 222-232, April 2000.
[4] Denning.D.E., Neumann.P.G, "*Requirements and Model for IDES- A Real-Time Intrusion Detection System*", *Technical Report, Computer Science Laboratory, SRI International,Menlo Park, California, pp.58-63,1985.*
[5] Gjorgji Madzarov, Dejan Gjorgjevikj and Ivan Chorbev,"*A Multiclass SVM Classifier Utilizing Binary Decision Tree*", *Informatica33, 2009, pp. 233-241.*
[6] Jun GUO, Norikazu Takahashi, Wenxin Hu, "*An Efficient Algorithm for Multiclass Support Vector Machines*", *IEEE-2008.*
[7] *KDD Cup 1999 Data, Information and Computer Science, University of California, Irvine.* http://kdd.ics.uci.edu/databases/kddcup99/kddcup99. html
[8] Lippmann R.P., Fried D.J., Graf I., Haines J.W., Kendall K.R., McClung D., Weber D., Webster S.E., Wyschogrod D., Cunningham R.K., and Zissman M.A.,"*Evaluating Intrusion Detection Systems: The 1998 DARPA Off-Line Intrusion Detection Evaluation*", *in Proceedings of the 2000 DARPA Information Survivability Conference and Exposition (DISCEX), Vol. 2, IEEE Computer Society Press: Los Alamitos, CA, pp.12-26, 2000.*
[9] Lee.W and S. J. Stolfo, "*A Framework for Constructing Features and Models for Intrusion Detection Systems*," *ACM Transactions on Information and System Security*, Vol. 3, No. 4, pp. 227-261, Nov. 2000.
[10] Latifur Khan, Momoun Award, Bhavani Thuraisingam," *A new Intrusion Detection System using Support Vector Machines and hierarchical clustering*", The VLDB Journal DOI 10, 1007/s00778 – 006 – 0002, 2007.
[11] Mahbod Tavallare.Y, Ebraham Bagheri, Wei Lu and Ali A.Ghortban,"*A Detailed Analysis of KDD Cup'99 data set*", *Symposium on Computational Intelligence Security on Defence Applications (CISDA),IEEE-2009.*
[12] Quinlan.J.R, "*C4.5: Programs for Machine Learning*", Morgan Kauffman, 1993.
[13] Sandya Peddabachigari, Ajth Abraham, Crina Grosan, Johnson Thomas, "*Modeling Intrusion Detection Systems using Hybrid Intelligent Systems*", *Journal of Network and Computer Applications-2005.*
[14] Stefan Axelsson, "*Intrusion Detection Systems: A Survey and Taxonomy*", Technical Report No 9, Dept. of Computer Engineering, Chalmers, University of Technology, Sweden, pp. 9-15, 2000.
[15] Snehal A.Mulay, P.R. Devale, G.V. Garje, "*Intrusion Detection System using Support Vector Machine and Decision Tree*", International Journal of Computer Applications, Volume3-No.3, pp.0975-8887, June 2010.
[16] Teresa L, Ann T and Fred G, "*A Real-Time Intrusion Detection Expert System(IDES)*" Technical Report, Computer Science Laboratory, SRI International, Menlo Park, California, pp.158-163,1992.
[17] Xiaodan Wang, Zhaohui Shi, Chongming Wu and Wei Wang, "*An Improved Algorithm for Decision Tree based SVM*", Intelligent Control and Automation, 2006. (WCICA 2006) , pp. 4234 - 4238 , IEEE -2006.
[18] Xiang.C, M.Y.Chong and H.L.Zhu, "*Design of Multiple-Level Tree Classifier for Intrusion Detection System*", In *Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems*, Singapore, pp.872-877, Dec. 2004
[19] Zhi-song Pan, Songcan Chen, Gen-bao Hu, Dao-qiang Zhang, "*Hybrid Neural Network and C4.5 for Misuse Detection*", Proceedings of the Second IEEE International Conference on Machine Learning and Cybernetics, pp. 2-5 , November 2003.
[20] Zhi-xin Yu, Jing-Ran Chen and Tian-Qing Zhu, "A Novel Adaptive Intrusion Detection System Based on Data Mining", In proceedings of the fourth international Conference on Machine learning and Cybernetics, Guangzhou, pp. 2390-2395, August 2005.
[21] Zeng, Q. and Wu, S. "Anomaly detection based on multi-attribute decision", WRI Global Congress on Intelligent Systems, Vol. 2, pp. 394-398, 2009.
[22] Kim, S.S. and Reddy, A.L.N. "Statistical techniques for detecting traffic anomalies through packet header data", IEEE/ACM Transactions on Networking, Vol. 16, No. 3, pp. 562-575, 2008.
[23] Cherkasova, L., Ozonat, K., Symons, J. and Smirni, E. "Automated anomaly detection and performance modelling of enterprise applications", ACM Transactions on Computer Systems, Vol. 27, No. 3, pp. 1-32, 2009.

An Efficient Illumination Normalization Method with Fuzzy LDA Feature Extractor for Face Recognition

Behzad Bozorgtabar[1], Hamed Azami[2]

(Department of Electrical Engineering / Iran University of Science and Technology, Iran)

ABSTRACT

The most significant practical challenge for face recognition is perhaps variability in lighting intensity. In this paper, we developed a face recognition which is insensitive to large variation in illumination. Normalization including two steps, first we used Histogram truncation as a pre-processing step and then we implemented Homomorphic filter. The main idea is that, achieving illumination invariance causes to simplify feature extraction module and increases recognition rate. Then we utilized Fuzzy Linear Discriminant Analysis (FLDA) in feature extraction stage which showed a good discriminating ability compared to other methods while classification is performed using three classification methods : Nearest Neighbour classifier , Support Vector Machines (SVM) and Feedforward Neural Network(FFNN).The experiments were performed on the ORL (Olivetti Research Laboratory) and Yale face image databases and the results show the present method with SVM classifier outweighs other techniques applied on the same database and reported in literature.

Keywords **- Face Recognition, Homomorphic filter, Nearest Neighbour, Fuzzy LDA, SVM, FFNN**

I. INTRODUCTION

Face recognition has become one of the most active research areas of pattern recognition since the early 1990s, and has attracted substantial research efforts from the areas of computer vision, bio-informatics and machine learning.

Illumination is considered one of the most difficult tasks for face recognition. The illumination setup in which recognition is performed is in most cases impractical to control, its physics difficult to accurately model and face appearance differences due to changing illumination are often larger than those differences between individuals. Reliable techniques for recognition under more extreme variations caused by pose, expression, occlusion or illumination is highly nonlinear, have proven elusive [1].

In this paper, we outline a hybrid technique for illumination normalization, after the Histogram truncation was applied to input face images, the Homomorphic filter was used for normalization. The face recognition involves two major steps. In the first step, some features of the image are extracted. In the second step, on the basis of the extracted features the classification is performed.

Fuzzy LDA (Fuzzy Fisherface) recently, was proposed for feature extraction and face recognition [2]. Fuzzy LDA computes fuzzy within-class scatter matrix and between-class scatter matrix by incorporating class membership of the binary labeled faces (patterns).

Finally extracted features were considered as inputs to classifiers. In this paper well-known classifiers including Nearest Neighbor, SVM and Feedforward Neural Networks were employed as classification.

Then rest of this paper is as followed. Our proposed method is initiated in second section then Fuzzy LDA is introduced in third section. Classification, experimental results on both ORL and Yale datasets, and Conclusion are depicted in continue.

II. ILLUMINATION NORMALIZATION TECHNIQUE

In this stage, in order to boost the result of normalization, we first truncated a specified percentage of the lower and upper ends of an image histogram.

In fact, several studies have shown that histogram remapping in conjunction with photometric normalization techniques results in better face recognition performance than using photometric normalization techniques on their own.

In the next step, Homomorphic filter as a renowned illumination reflectance was used. And then filtered face image is considered as input of feature extraction module.

A. Homomorphic Filter
Homomorphic filtering (HOMO) is a well known normalization technique, which improves the

pearance of an image by contrast enhancement and gray-level range compression.

Consider an image, $f(x, y)$, which can be stated as the product of the illumination $i(x, y)$, and the reflectance component $r(x, y)$ as follows [3]:

$$f(x, y) = i(x, y). \, r(x, y) \qquad (1)$$

Then input image is transformed in to the logarithm domain in order to achieve frequency components of the illumination and reflectance separately:

$$z(x, y) = \ln f(x, y) \qquad (2)$$

$$= \ln i(x, \, y) + \ln r(x, y)$$

Then:

$$\{ \, z(x, \, y) \, \} = \mathfrak{F} \, \{ \, \ln f(x, \, y) \}$$

$$= \mathfrak{F} \, \{ \, \ln i(x, \, y) \} + \, \mathfrak{F} \{ \, \ln r(x, \, y) \}$$

Or:

$$Z(u, \, v) = F_i(u, \, v) + F_r(u, \, v)$$

Where $F_i(u, \, v)$ and $F_r(u, \, v)$, in equation (2) are the Fourier transforms of the term defined.

The Fourier transform of the product of the $Z(u, v)$ and filter function $H(u, v)$ can be expressed as:

$$S(u, \, v) = H(u, \, v). \, Z(u, \, v) \qquad (3)$$

$$= H(u, \, v). \, F_i(u, \, v) + H(u, \, v). \, F_r(u, \, v)$$

In the spatial domain:

$$s(x, \, y) = \mathfrak{F}^{-1}\{ \, S(u, \, v)\}$$

$$= \mathfrak{F}^{-1}\{H(u, \, v). \, F_i(u, \, v) + H(u, \, v). \, F_r(u, \, v)\}$$

Finally by letting

$$i(x, \, y) = \mathfrak{F}^{-1}\{ \, H(u, \, v). \, F_i(u, \, v) \, \} \qquad (4)$$

$$(x, \, y) = \mathfrak{F}^{-1}\{ \, H(u, \, v). \, F_r(u, \, v) \, \}$$

the equation becomes:

$$s(x, \, y) = i(x, \, y) + r(x, \, y) \qquad (5)$$

$$g(x, \, y) = e^{\, s(x,y)}$$

$$g(x, \, y) = e^{\, i(x,y)} + e^{\, r(x,y)}$$

$$g(x, \, y) = i_0(x,y) + r_0(x, \, y)$$

Where i_0 and r_0 are the illumination and the reflectance components of the output images. After $z(x, \, y)$ is transformed into the frequency domain, the high frequency components are emphasized and the low-frequency components are reduced. As a final step the image is transformed back into the spatial domain by applying the inverse Fourier transform and taking the exponential of the result.

This method is based on a special case of a class of systems known as Homomorphic system. The filter transform function $H(u, \, v)$ is known as the Homomorphic filter[4] .

III. FUZZY LDA (FLDA)

Fuzzy LDA, which also was called Fuzzy Fisher Face method, is considered to solve binary classification problems. In conventional LDA approach, every vector is supposed to have a crisp membership. But this does not take into account the resemblance of images belonging to different classes, which occurs under varying conditions. In FLDA, each vector is assigned the membership grades of every class based upon the class label of its k nearest neighbours. This Fuzzy k-nearest neighbour is utilized to evaluate the membership grades of all the vectors [5].

The membership degree to class i for j^{th} pattern is obtained from following equation[5]:

$$u_{ij} = \begin{cases} 0.51+0.49\left(n_{ij}/k\right) & i, j \, belong \quad to \quad the \quad same \quad class \\ 0.49\left(n_{ij}/k\right) & otherwise \end{cases} \qquad (6)$$

In the above expression n_{ij} stands for the number of the neighbors of the jth data (pattern) that belong to the ith class. As usual, u_{ij} satisfies two obvious properties:

$$\sum_{i=1}^{C} u_{ij} = 1 \qquad (7)$$

$$0 < \sum_{j=1}^{N} u_{ij} < N \qquad (8)$$

Therefore, the fuzzy membership matrix U can be achieved with result of FKNN.

$$U = [u_{ij}] \, i = 1, 2, ..., c \quad j = 1, 2, ..., N \qquad (9)$$

The results of the FKNN are used in the computations of the statistical properties of the patterns.

Taking in to account the fuzzy membership degree, the mean vector of each class is [6]:

$$m_i = \frac{\sum_{j=1}^{N} u_{ij} x_j}{\sum_{j=1}^{N} u_{ij}} \qquad (10)$$

Then, the membership degree of each sample (contribution to each class) should be considered and the corresponding fuzzy within-class scatter matrix and fuzzy between-class scatter matrix can be redefined as follow [7]:

$$FS_W = \sum_{i=1}^{c} \sum_{x_j \in w_i} u_{ij}(x_j - m_i)(x_j - m_i)^T \qquad (11)$$

$$FS_b = \sum_{i=1}^{c} \sum_{j=1_i}^{N} u_{ij}(m_i - \bar{X})(m_i - \bar{X})^T \qquad (12)$$

Where $\bar{X}$, is the mean of all samples. So, all scatter matrices with fuzzy set theory are redefined and the contribution of each sample is incorporated.

Our optimal fuzzy projection W_{F-LDA} follows the expression:

$$W_{F-LDA} = \arg\max_w \frac{|W^T FS_B W|}{|W^T FS_W W|} \qquad (13)$$

It is difficult to directly calculate W_{F-LDA} because that FS_W is often singular [6]. For tackle this problem, PCA is used as a dimension reduction step and thus the final transformation is given by the following matrix,

$$W^T = W_{F-LDA}{}^T W_{PCA}{}^T \qquad (14)$$

IV. CLASSIFICATION
A. Nearest Neighbour Classifier
After evaluating feature vectors by FLDA, test images are projected on feature space and the distances to the training images are computed using nearest neighbour algorithm for the purpose of classification as follows [8]:

For two images i and j, let $f^{(i)}$ and $f^{(j)}$ representing the corresponding feature vectors, the distance d_{ij} between the two patterns in the feature space is defined as:

$$d_{ij} = \sqrt{\sum_n \left(\frac{f_{(n)}^{(i)} - f_{(n)}^{(j)}}{\alpha(f_n)} \right)^2} \qquad (15)$$

Where $f_n^{(i)}$ is the *nth* element of the feature vector i while the term $\alpha(f_n)$ is the standard deviation of the *nth* element over the entire database and is used to normalize the individual feature components. Finally, a test image j is assigned to image i in a database with the smallest corresponding distance d_{ij} .

B. Feedforward Neural Networks (FFNN)
FFNN is suitable structure for nonlinear separable input data. In FFNN model the neurons are organized in the form of layers. The neurons in a layer get input from the previous layer and feed their output to the next layer. In this type of networks connections to the neurons in the same or previous layers are not permitted. Fig 1 shows the architecture of the system for face classification [9-10].
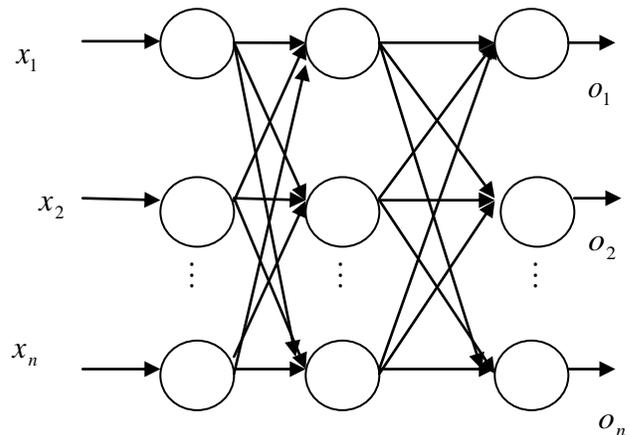
Fig.1. Architecture of FKNN for classification

In this experiment the number of nodes in hidden layer is set to 15. A large neural network for all people in the database was implemented. After calculating the features, the feature projection vectors are calculated for the faces in the database. These feature projection vectors are used as inputs to train the neural network. Fig 2 illustrates the schematic diagram for the NN training phase.

C. Support Vector Machine (SVM)
SVM is a binary classification method that intends to find the optimal linear/nonlinear decision surface based on the concept of structural risk minimization. The decision surface is a weighted representation of the elements of training set. The elements on the

decision surface are defined by a set of support vectors which characterizes the boundary between two (or more) classes [11]. The problem of multi-class is solved by combining multiple two class SVMs. We select the one-versus-the-rest approach that constructs SVMs which the train k^{th} model chooses the kth class as the positive examples and the remaining $(k-1)$ classes as the negative examples. Comparison with one-versus-one, it significantly needs less training time. At last, indexed SVM classifier is learned through quadratic programming in order to find the class's boundaries with maximum margins. This technique helps to find the accurate relations between nearest similarity faces.

V. SIMULATION AND RESULTS

A. ORL Face Database

The ORL database consists of 40 groups, each containing ten 112×92 gray scale images of a single subject [12]. Each subject's images differ in lighting, facial expression, details (i.e. glasses/no glasses) and even sliding. Some of the database's images are illustrated in Fig 3.

The ORL database consists of 40 groups, each containing ten 112×92 gray scale images of a single subject. Each subject's images differ in lighting, facial expression, details (i.e. glasses/no glasses) and even sliding. Some of the database's images are illustrated in Fig 3.

Fig 4 shows an original image from ORL database and its histogram respectively. This image is chosen specifically because it had led to misclassification in many feature extraction methods. Being taken under ambient lighting in a neutral facial expression and the person wore glasses, the images of this class lead to increase in error rate.

In the next step,the Homomorphic filter was applied to the selected image for normalization. The filtered image and its histogram are displayed by Fig5.

Initially we implemented the Histogram truncation, in this step the lower and upper ends of an image histogram that must be truncated, were set to 20 percent and 60 percent respectively.

Then Homomorphic filter was used, the performance of this filter rely on its parameters. We set the cut-off frequency of the filter to 0.5 and second order of the modified Butterworth style filter is used. In the next step, we shortened half of upper ends of final histogram.

After preprocessing step, FLDA was applied in order to extract features. Then described classifiers are implemented.

Fig.2. Training stage for Neural Network

Fig.3. Samples of ORL face database
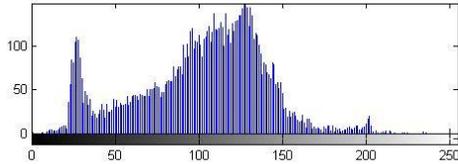
Fig.4.Original image and the corresponding histogram
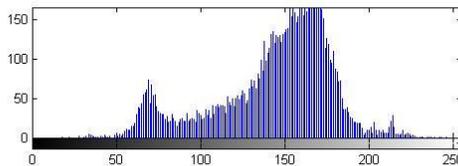
Fig.5.Filtered image and the corresponding histogram

Two set of images were created from the ORL face database; For the Five-to-Five dataset, five random images of each group were selected for training while the others were used for testing. For the Leave-One-Out set, 9 images were used for training and the remaining image was kept for validation.

Table 1 shows comparison among different methods in Five-to Five set that simulation results show SVM classifier with proposed method has a better performance compared to other classifiers.

Table 1. Comparison of recognition rates of different classifiers based on various feature extractors for ORL face database

	Nearst Neighbour	SVM	FFNN
PCA	88%	90%	89%
LDA	90%	92%	91%
FLDA	94%	95%	92%
2D-PCA	93%	94.5%	93%
Proposed	95%	96.5%	94.5%

Table2 illustrates recognition rates of depicted methods with nearest neighbour classifier for two sets.

Table2. Comparison of recognition rates of different methods for two sets

	Five-to-Five	Leave one -out
PCA	88%	89.%
LDA	90%	91%
FLDA	94%	95%
2DPCA	93%	94.%
Proposed	95%	96%

B. Yale Face Database

The Yale face database contains 165 images of 15 individuals (each person providing 11 different images) under various facial expressions and lighting conditions [13]. Fig 6 shows sample images of one person.

Similar procedure was done for Yale face database. In this set 6 images were used for training from each class and remaining images were employed for test module. As far as recognition rate is concerned proposed method outranks others with 96.5 and 95.5 percent for SVM and two other classifiers respectively. In addition the outcomes of Feedforward Neural Networks and Nearest Neighbour are close to each other. The results were shown in Table 3.

Fig.6. Samples of Yale face database

Table3. Comparison of recognition rates of different classifiers based on various feature extractors for Yale face database

	Nearest Neighbour	SVM	FFNN
PCA	89%	90%	90%
LDA	91%	93%	92%
FLDA	93.5%	94%	91%
2D-PCA	94%	95%	93.5%
Proposed	95.5%	96%	95.5%

VI. CONCLUSION

In this paper, the Face Recognition problem was addressed by improved method based on modified Homomorphic filter, which is insensitive to large variation in illumination.

Homomorphic filter is a celebrated normalization technique which was ignored in face recognition. In continue Fuzzy LDA (FLDA) was used that compared to other feature extractors, it had a good ability in discriminating of classes.

Finally proposed method in collaboration with appropriate classifiers was used which the result showed proposed method with SVM classifier had a best performance.

REFERENCES

[1] Y A. Georghiades, P. Belhumeur, and D. Kriegman,, From few to many: Illumination cone models for Face Recognition under variable lighting and pose, *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6),* 2001, 643-660.

[2] Kw K C, Pedry W, Face recognition using a fuzzy fisher classifier, *Pattern Recognition, 38(10),* 2005, 1717-1732.

[3] Thammizharasi,A.M.E, Performance Analysis Of Face Recognition By Combining Multiscale Techniques And Homomorphic filter using Fuzzy k-nearest neighbour classifier, *Proc. IEEE Conf. on Communication Control and Computing Technologies (ICCCCI'10)*, 2010, 643-650.

[4] N. Ahmed Surobhi , Md. Ruhul Amin, Employment of Modified Homomorphic filters in Medical Imaging, *International University Journal of science and Technology in Daffodil, 1(1),* 2006.

[5] M.Keller, M.R.Gray, J.A.Givern, A Fuzzy K Nearest Neighbor Classifier Algorithm, *IEEE Transactions on Systems, Man and Cybernetics, 15(4),* 1985,580-585.

[6] W. Yang, Hui Yan, J. Wang, J. Yang, Face Recognition using Complete Fuzzy LDA, *Proc. 19th International Conf. on Pattern Recognition*, ICPR,2008, 1-4.

[7] X Song,Y Zheng, A complete Fuzzy Discriminant Analysis Approach for Face Recognition, *Applied Soft Computing,* 2010,208-214.

[8] Peter N. Belhumeur, Joao P. Hespanha, David J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7),* 1997.

[9] A. Eleyan , H. Demirel, Face recognition system Based on PCA and Feedforward Neural Networks, *Proc. Computational Intelligence and Bioinspired Systems*, Barcelona, Spain, 2005, 935-942.

[10] Lawrence, S., Giles, C. L., Tsoi, A. C., Back, A. D., Face Recognition: A Convolutional Neural-Network Approach, *IEEE Transactions on Neural Networks, 8(1),* 1997.

[11] J. Huang, V. Blanz and B. Heisele, Face recognition using component-based SVM. Classification and morphable models, *Lecture Notes in Computer Science, 23(88),* 2002, 334-341.

[12] ORL, "The Database of Faces," 2011. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatase.html.

[13] Yale, "The Database of Faces," 2011. cvc.**yale**.edu/projects/**yalefaces**/**yalefaces**.html

Optimization of a water cooled Submersible motor journal bearing length using CFD

Dr. Virajit A. Gundale[1], Mr. Sanjaykumar M. Ingale[2], Vidyadhar M. Dandge[3]

[1]Professor & Academic-Incharge, Sharad Institute of Technology College of Engineering, Yadrav Dist. Kolhapur India.

[2]Associate Professor & Head, Department of Mechanical Engineering, Sou. Sushila Danchand Ghodawat Charitable Trust's Group of Institutions, Post-Atigre, Tal.–Hatkanangale, Dist.– Kolhapur

[3]Asisstant Professor, Dr.J.J.Magdum College Of Engineering Jaysingpur, Dist-Kolhapur, India

ABSTRACT

The journal in case of a water cooled Submersible motor is the bearing surface of the rotor which consists of hard chromed stainless steel material. The bearing consists of a Leaded bronze material (Bronze grade LTB 2,3 or 4 of IS 318 or Nitrile / cutless rubber) which is the softer part out these two. The failures of the bearing bush accounts for 90% of the failures in a submersible motor. The implication of this failure is the complete breakdown of the Motor i.e. the windings will get damaged and expensive repairs would be required to be carried out. This clearly suggests that the design of such a journal should be properly investigated. A CFD approach would assist in establishing the dimension (length) of the bearing. This paper presents step by step application of CFD to optimize the bearing length and is an outcome of around 3 years of extensive research in an attempt to solve a manufacturer's long standing problem.

Keywords- Submersible motor, journal, bearing, rotor, CFD

I. INTRODUCTION

In India the water cooled type Submersible Motors are extensively manufactured and available in the market due to its simplicity in design and manufacture. The maintenance of such motors is also very simple and can be carried out at ease compared to the oil filled version.

Even with such advantages water cooled submersible motors too pose various problems especially with its bearing bushes as indicated in Figure 1.0 and 1.1.

Figure 1.0 Worn out bearing bush (Courtesy VIRA PUMPS)

Figure 1.1 Worn out bearing surface of the Rotor, (Courtesy VIRA PUMPS)

II. PROBLEM IDENTIFICATION

This was a long standing problem at M/s VIRA PUMPS, Kolhapur, Maharashtra, INDIA for around 5 years. This Industry is a reputed manufacturer and exporter of Submersible Pumps. It has started producing 100 mm (4") Submersible motors since 2001. Figure 1.2 shows a sectional view of such a Submersible motor
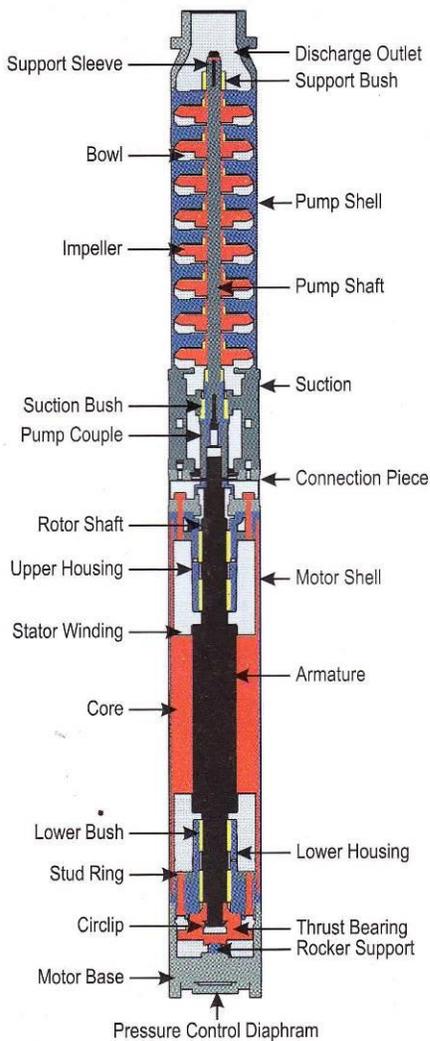
Figure 1.2 Sectional View of a 100 mm (4") Submersible motor

For motors above 1.5 hp, it experienced bearing bush failures after just few months of operation. Where as its earlier products i.e. 150 mm and 200 mm Submersible motors operated smoothly for more than 25 years. Due to this the Industry faced huge problems in their operations. Their reputation had been at stake. They were now thinking to discontinue this problematic range of products. They were not only the ones who suffered but, similar manufacturers in India experienced the same problem.

        This problem was taken as a challenging project in 2006 as an attempt to save this particular range of Motors and to recover the loss incurred for the last few years. An extensive study was made by referring various literatures as well as the IS guidelines for manufacturing this particular motors. Several national and international

brands of these specific motors were thoroughly analyzed.

The root cause analysis suggests that the following factors are responsible for such bearing bush failures:

1. Material of the bearing bush.
2. Poor Surface finish of the journal.
3. Wrong length and clearances maintained.
4. Machining defects like run-outs, etc.
5. Overall design of the Motor.

The materials being used absolutely confirmed to IS 318 with correct percentage of copper and minimum lead as the basic requisites. So, the material issue was ruled out. The surface finish was also maintained below 0.7 microns which was desired in the range of 0.2 to 0.4 microns. There was also no problem whatsoever in the run outs, etc. as the rotor was dynamically balanced on superior balancing machine. This came to the decision that the length of the Bearing bush needed to be investigated.

        As a test the bearing length was increased up to 30% and the said motor being manufactured and assembled. The motor was coupled with a suitable pump and the system was installed and run for around a year.

        The results were encouraging as the bearing bush did not failed at all. This pointed out that the long standing problem can be solved by evaluating the correct bearing length and optimizing the same scientifically. The important factor i.e. the Permissible bearing pressure is vital in the design of such bearings. This calculation can be assisted with a CFD analysis to establish its correct value of this factor so as to assist correct calculations.

III. CFD ANALYSIS OF JOURNAL BEARING :
The conventional method in designing a Journal bearing is by using a bearing pressure recommended for specific application. In the case of a Submersible Motor it is recommended to use a bearing pressure in the range 0.7 to 1.4 as shown in the table 1.0

Machinery	Bearing	l/d	Permissible bearing Pressure (N/mm^2)
Gas and oil engines (4-stroke)	Main Crank pin Wrist pin	0.6-2.0 0.6-1.5 1.5-2.0	4.9-8.4 10.8-12.6 12.5-15.4
Gas and oil engines (2-stroke)	Main Crank pin Wrist pin	0.8-1.8 0.7-1.4 1.5-2.2	5.6-11.9 10.5-24.5 16.1-35.0

Aircraft and automobile engines	Main	0.8-1.8	5.6-11.9
	Crank pin	0.7-1.4	10.5-24.5
	Wrist pin	1.5-2.2	16.1-35.0
Reciprocating compressors and pumps	Main	1.0-2.2	1.75
	Crank pin	0.9-1.7	4.2
	Wrist pin	1.5-2.0	7.0
Centrifugal pumps, motors and generators	Rotor	1.0-2.0	0.7-2.0
Railway cars	Axle	1.9	3.5
Marine steam engines	Main	0.7-1.5	3.5
	Crank pin	0.7-1.2	4.2
	Wrist pin	1.2-1.7	10.5
Punching and shearing machines	Main	1.0-2.0	28
	Crank pin	1.0-2.0	56
Rolling Mills	Main	1.0-1.5	21

Table 1.0 Permissible bearing pressure

Instead of taking the value directly from the above table, we will perform CFD Analysis on the bearing. By this we will also be able to verify the value of bearing pressure 'p'. We can then use the value obtained from the analysis and perform the design steps to calculate the length of bearing.

Figure 1.5 shows the assembly for the CFD Analysis. Solidworks Flow simulation software is used for the CFD Analysis. Figure 1.6 represents the Surface plot indicating the Maximum bearing pressure. The journal in this case is rotated at 2800 rpm which is the rated speed of the motor. The centers of the journal and bearing bush are offset to around 0.015 to mimic the real working condition. The working fluid is chosen as water.
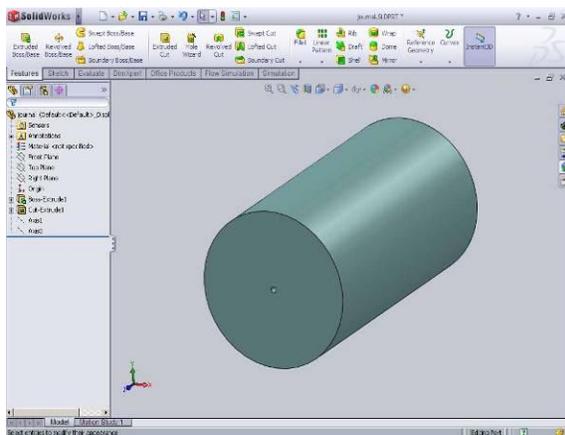
Figure 1.3 Representation of the journal
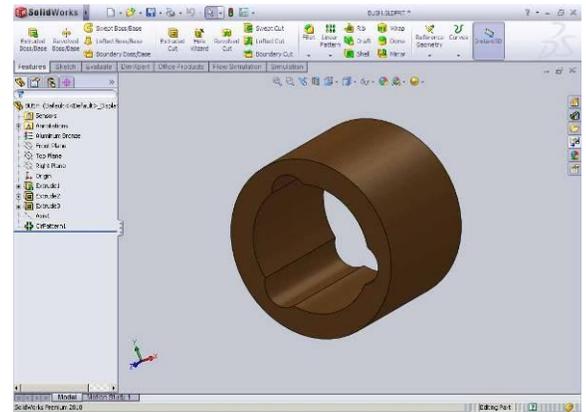
Figure 1.4 Bearing bush of the Submersible Motor
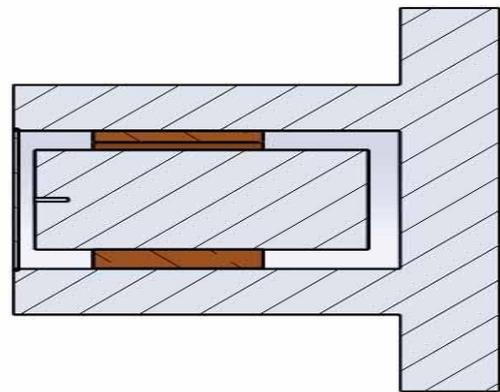
Figure 1.5 Assembly for the CFD analysis
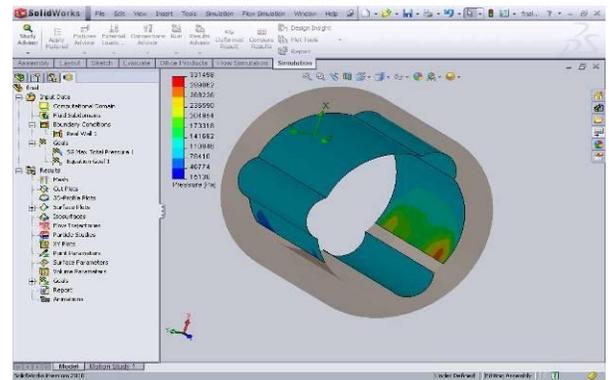
Figure 1.6 Surface plot indicating Bearing pressure

So, from this analysis we get the bearing pressure p=0.38 or 0.4 N/mm². We will use this value in the design of our journal bearing as follows:

We will consider the case for the design of the Journal bearing for a 1.5 hp or 1.1 kW submersible motor which rotates with a constant speed of 2800 rpm or 293 rad/s. We need to find the Radial load F_N which is given by,

$$F_N = \frac{9550 \times kW}{N \times R} \qquad \ldots(1)$$

Where R is the radius of the journal. In this case it is

0.01299 m

$$= \frac{9550 \, x \, 1.1}{2800 \, X \, 0.01299}$$

= 288.82 N

Now, Area A = l x d

= l x 25.98

We already know that,

p=w/ld   …(2)

taking value of p= 0.4 N/mm^2 and

w= F_N=288.82 N

so,

$$0.4 = \frac{288.82}{l \, x \, 25.98} \qquad …(3)$$

Thus, l= 27.79 mm which is the bearing length. At present the bearing length used for 1.5 hp or 1.1 Kw motor is 24 mm. The recommendation made here is that the bearing length must be changed to 27.79 mm.

Let us check this by using a bearing pressure of 0.7 N/mm^2 is used as per Table 1.0

From (3),

$$0.7 = \frac{288.82}{l \, x \, 25.98}$$

The length comes out to be 15.9 mm which is very less. Again if the value of bearing pressure is increased say in steps till 1.4 then it is quite obvious that we will get an in correct length of the bearing bush.

IV. RESULTS AND DISCUSSIONS

The specified bearing pressure for this particular application was 0.7 N/mm^2 to 1.4 N/mm^2 as per Table 1.0 above. Using these values it is found that the bearing lengths obtained are very short and may be not sufficient for that specific power rating. Thus, a precise bearing pressure has to be used to get accurate results. It has been found that by implementing the bearing pressure obtained by CFD results gives a more accurate length of the bearing. The same technique can be used for various diameters of the journal or the bearing bush and for various power ratings which ultimately help in achieving an optimum journal bearing length can be achieved.

The manufacturer is an ISO 9001 Certified organization. The basic objective of the organization is to control rejection below 3 %. The new design was implemented in October 2009. The results were collected in May 2011 i.e. after 17 months as shown in figure 1.7
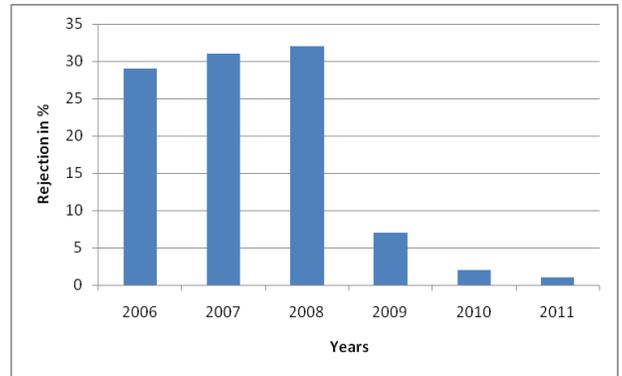
Figure 1.7 Rejection percentage

This new method and design changes were communicated to all potential customers and practical demonstrations of the pump sets were done in presence of customers. This boosted confidence of the customer that the failure reasons have been removed and product is updated. Slowly the customer response gone up and we observed stiff rise in quantum of sales of this product which is graphically represented in Figure 1.8 This is very much fruitful achievement of the research undertaken and completed which was related to Bearing bush of the Submersible motor.
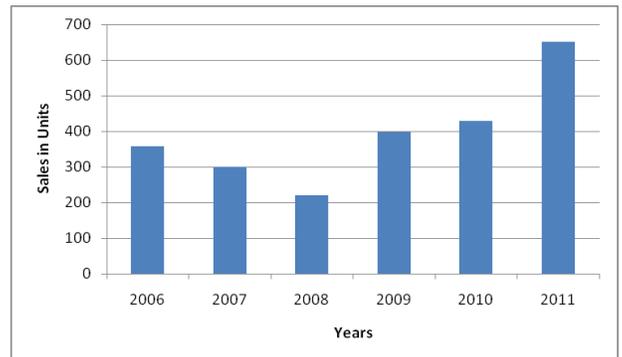
Figure 1.8 Increase in Sales

V. CONCLUSION:

This paper thus demonstrates how CFD Technique can be effectively used to sort out design lacunas in existing designs and make the product free from complaints. This saves sizable amounts of effort and money. The manufacturing organizations who are engaged in design and development of own products must establish a system which will keep consistent focus on field operation of the product through immediate corrective and preventive actions to improve upon the product life cycle through harnessing available advanced and most suitable technologies.

ACKNOWLEDGEMENTS

REFERENCES:

[1] Sternlicht, B., Hydrodynamic Lubrication, Mechanical Design and Systems Handbook, Rothbart, H.A. (Ed.), 2nd ed., McGraw Hill, N.Y. (1986).

[2] P. Kannaiah *Machine Design*, Scitech Publications (India) Pvt.Ltd., 2009,Second Edition.

[3] Joseph E. Shigley, Charles R. Mischke *Mechanical Engineering Design*, Tata Mcgraw-Hill Publishing Company Ltd.,2001 Sixth Edition.

[4] Versteeg, H.K. and Malalasekera, W., 2007, An introduction to Computational Fluid Dynamics: the Finite-Volume Method (2nd edition), Pearson.

[5] Ferziger, J.H. and Peric, M., 2001, *Computational Methods for Fluid Dynamics (3rd edition)*, Springer-Verlag.

[6] IS 9283 : 1995, Motors for Submersible Pumpsets-Specification (First Revision).

[7] Virajit Avinash Gundale, 2010, *A new design approach for water cooled submersible motor and radial flow type pump with emphasis on both Electrical and Mechanical consideration'* PhD Thesis, UNEM, Costa Rica.

ABOUT THE AUTHORS

Dr. Virajit A. Gundale is presently working as the Professor & Academic In-charge at Sharad Institute of Technology College of Engineering, Yadrav Dist. Kolhapur, India. Apart from this he is well known consultant in the design and development of Submersible pumps and motor components. He has worked on various international projects in Bangladesh, Indonesia, Egypt, etc. He obtained his B.E. in Mechanical Engineering from Shivaji University and M.S. in Manufacturing Management from BITS, Pilani. He obtained his Ph.D. in Manufacturing Technology from UNEM, Costa Rica in 2010. His total experience including Teaching and Industry spans more than 12 years. He is also a fellow of the International Institute of Mechanical Engineers, South Africa.

Sanjaykumar M. Ingale is presently working as the Associate Professor at Sou. Sushila Danchand Ghodawat Charitable Trust's Group of Institutions, Post-Atigre, Tal.–Hatkanangale, Dist-Kolhapur, India. He obtained his B.E. in Production Engineering from Shivaji University and M.E (Production) from Shivaji University. His total experience including Teaching and Industry spans more than 10 years.

Vidyadhar M. Dandge is presently working as the Assistant Professor at Dr.J.J.Magdum College Of Engineering Jaysingpur, Dist-Kolhapur, India. He obtained his B.E. in Mechanical Engineering from Karnataka University and M.E (Production) from Shivaji University. His total experience including Teaching and Industry spans more than 25 years.

DESIGN AND ANALYSIS OF AMLA PUNCHING MACHINE

Prof .R.B.Salwe
Asst. Professor, Mechanical Engg.
Bapurao Deshmukha College of
Engineering Sevagram , India

Prof. S.M.Fulmali
Asst. Professor, Mechanical Engg.
Bapurao Deshmukha College of
Engineering Sevagram , India

Prof. S.B.Khedkar
Asst Professor, Mechanical Engg.
Bapurao Deshmukha College of
Engineering Sevagram , India

Prof. S.Y.Sonye
Asst Professor, Mechanical Engg.
Bapurao Deshmukha College of
Engineering Sevagram , India

Abstract—
In India, amlas are produced on a large scale which is used for various purposes such as making murraba, pickles, etc. But the method of making murabba by manual method is less efficient and not suitable to maker. The present work is about implementation of 3D tools in the optimum utilization of amla by punching it in appropriate manner.
In rural areas, such type of machine is very useful and demanded. The making of amla's murabba by manual method consumes more time. This machine is not yet manufactured or invented, so the present work is to design the machine which will reduce the time and increase efficiency. Also many industries (such as Agrosaw pvt limited) are working on this concept. The present work will be utilize small scale to medium scale food processing industries.

Keywords- Design; Machine; Punching;

I. INTRODUCTION

The Amla or Aonla (EmbJica OffiCinalis Gaertn) also known as Indian Gooseberry is a minor sub-tropical deciduous tree indigenous to Indian sub-continent and it can be grown successfully in dry and neglected regions owing to its hardy nature, suitability to various kinds of wasteland.

Amla has been in use for pickle and preserve since ages in India and the methods employed were based on traditional knowledge of grandmothers. Besides, amla has been an important

Ingredient for chavanprash, aayurvedic health tonics. The methods used previously were unhygienic in nature and time consuming. The nutrient loss in these methods was higher.

The current paper describes the use of amla for making murabba. The present work describes the use of manual method to punch the amla for making murabba .But, the problem is that, minor accidents have also been reported during manual seed removing, pricking and shredding. The shelf life of the manually prepared products was also less and the quality not up to the mark. So, that the amla punching machine concept was generated, The present work is about designing an machine to make it affordable for small scale industry .Till now most of the small scale industries used to punch amlas manually which can prove harmful to the workers. This process is much more time consuming, has less productivity.

Specially in rural areas most of the people does household business or many organizations in rural areas such as Mahila Bachat Gat does this business of punching the amlas and all the operations done there are manual which result in less production and thus less earning.

This paper shows that machine has the capability to overcome this problem and to make life better for the people doing this business in rural areas.

Objective to Design Machine

- Low cost
- High productivity
- Less harmful
- Less power consumption
- No supervision required
- Less man power required

II.   PARTS SELECTION FOR MACNHINE

FRAME:- THE FRAME IS A RIGID STRUCTURE MADE UP OF STRUCTURAL STEEL. IT CONSISTS OF THREE CHAMBERS. THE TOP CHAMBER CONSISTS OF HOPPER. THE MIDDLE ONE IS USED FOR THE PAIR OF CYLINDRICAL ROLLING DRUMS, SHAFT AND BEARINGS. THE BOTTOM CHAMBER IS USED FOR PLACING COLLECTOR. AT THE BASE, MOTOR IS PLACED TO RUN THE MACHINE TO GET THE DESIRED RESULT. THE DIMENSIONS OF THE FRAME ARE 87.2x40x5.

It is made up of thick Structural steel sheet. It is located at the top of the frame; at the middle of the hopper cam shape device is located which allows the amla to flow in two or more rows sequentially.


Figure 1. Frame of Punching Machine   Figure 2. Hopper
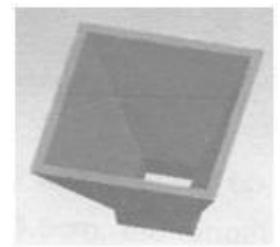
The dimensions of the hopper are 34x34x2S and the thickness is of 2 cm. It is shown in Figure 1 and 2.

Cylindrical Drum: - It is of cylindrical in shape and is made up of structural. Diameter of drum is between 3.5cm. The length is of 25cm.the circumference of the drum the pointed needles are mounted which is used for punching the amlas. Drum is mounted the shaft which is driven by drives. It is located at the middle of machine. It is shown in figure 3.

Figure.3. Cylindrical Drums

Punching Needles :- It is made up of stainless steel which is located on the circumference of the drum. The needles are specially fabricated for which the screws are used to join the pointed needles with drum.Total number of needles on circumference of drum is 240. It is shown in figure 4.
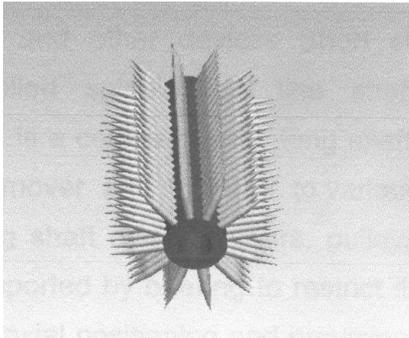

Figure 4.   Punching Needle

Bearings :- Rolling element bearing involve separation of the shaft and outer member by balls or rollers and thus substitute rolling friction by sliding friction. Since the contact areas are small and the stresses high, loaded parts of rolling element bearings are normally made up of hard, high strength material, superior to those of the shaft and outer member. These parts include inner and outer rings and the balls or rollers.

A major advantage of rolling element bearing is low starting friction. Rolling bearings are ideally suited for application involving high starting loads. Rolling element bearing take up more radial space around the shaft, but fluid film bearings required greater axial space. Rolling element bearing generate and transmit a certain amount of noise whereas fluid film earing do not normally generate noise and may dampen noise from other sources. Another advantage of rolling element bearing is that they can be preloaded. This is important in application requiring precise positioning of the rotating member. The rolling element bearing are also called anti friction a bearing which is a misnomer because this bearing does not always provide a lower friction than fluid film bearings.

Flat Belt:- The flat belt consists of strong elastic core surrounded by chrome leather or rubber. The elastic core consists of number of thin plies, made up of cotton, rayon, or nylon. For special application textile belts or balata belts are used. Flat belts are very efficient for high speed, they are quite, they can transmit large power over long centre distances, they don't require large pulley and they can transmit power around corner or between pulleys at right angle to each other.

Pulleys :- A pulley can lift very large masses a short distance. To illustrate the way it works in our machine, both pulleys at the top are of the same size, a long pulley on the belt will raise the section of belt extending down to the lowermost pulley by a similar length to the amount of belt lowered by the other top pulley. The weight will therefore go nowhere. The diameter of the pulley is 4.7cm.

Amla removal tool: - There are various type of amla removal tool .The amla removal tool is made up of rubber pad or rubber fiber it is placed below the each pair of drum. It will work only when the amla get stuck in to the needle.

Safety Quard: - It's made up of thin aluminum sheet. It is used to keep the amla between two drums. As there is possibility of bouncing and misplacing of amlas at the time of falling from the hopper. It also reduces the manual accident.

Collector: - It is a tray which is made up of aluminum sheet and in the lower portion of the net is provided. It is fixed inclined at the base of the frame.

Working of Machine:-In this machine, the cylindrical needle platform punches the amla's on its whole circumference. A single amla is punched more than ten times on its whole surface .The depth of the punches are 10-15 mm. After getting punched, the amlas are discharged from the machine and can be connected on plastic cranes. The machine is complete with electric motor, starter, etc.Amla punching machine is simple machine which is used for punching the amla in proper way to prepare   morraba.
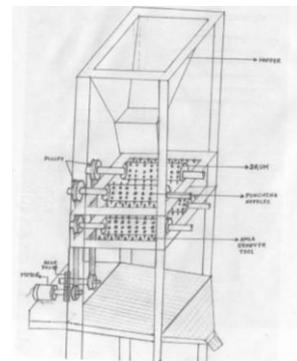
Figure 5.Line diagram of Punching Machine

Assembly of Machine:-
  Computer Aided Assembly of amla punching machining and bill of machine is shown in figure 9 and 10.
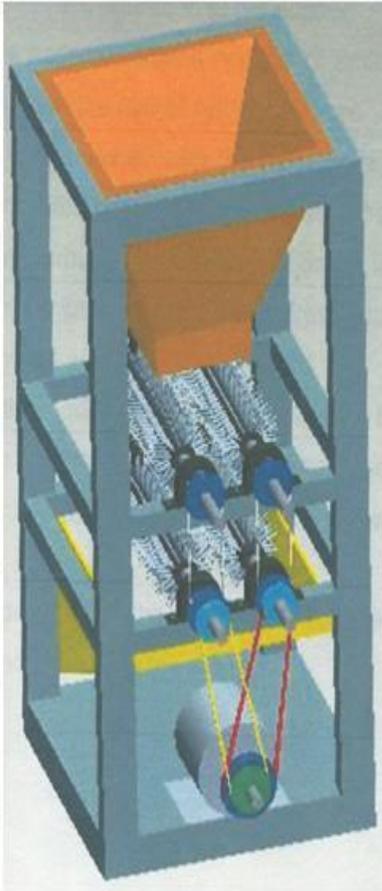
Fig.9.Model of punching machine
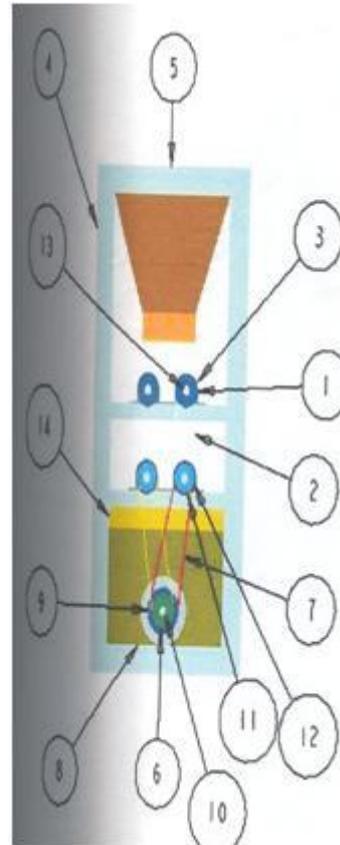
Fig.10.Bill of Machine Components

This machine works in the following way.

• First select the suitable numbers of amlas for punching.
• Selected amlas are poured in to the hopper.
• Through the hopper, suitable numbers of amlas are dropped below in between the pair of cylindrical drum.
• Amlas are then punched with the help of punching needles which are mounted on the circumference of the drum.
• After passing through the first pair of drum, then amlas are dropped on the second pair of cylindrical drum so that the amlas are punch again.
• In this way the numbers of holes are made on the amlas.
• If the amla are hanged in the punching needles. A special purpose amla removal tool is used to remove the hanged amlas.
• At last the punched amlas are collected in the collector and then it is used to prepare the morraba.

III.   FINITE ELEMENT ANAYSIS OF MACHINE

FEA with differential equation of system and end with solving them approximately. It goes through a number of steps in between. It converts differential equation in to integral equation by using variation approach or weighted residual method. Next it divides the problem domain into elements and develops the elements equations. It assembles the element equation to obtain the global system matrix equations. The boundary condition and external loads are applied to this system before solving. The result of the solution are available at the nodes of the elements .finite element analysis can display them in graphical form to analyse them, to make design decisions and recommendations.Conventional analytical method for solving stress and strain becomes very complex. In such cases finite element modeling becomes very convenient means to carry out the analysis. Finite element process allow for discretizing the intricate geometries into small fundamental volumes called finite element. It is possible to write the governing equations and material properties for these elements. These elements are then assembled by taking proper care of constraints and loading, which result in set of equations .these equations when solved give the result that described the behavior of original complex body being analyzed. Mesh model of machine is shown in figure 11.
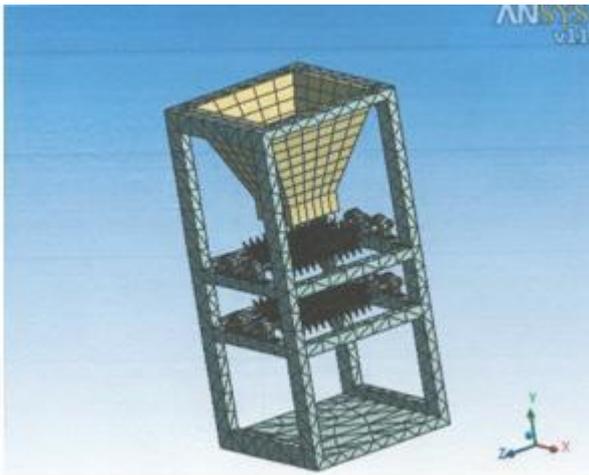
Figure 11. Discretization of Machine

Total no of nodes 118716 and elements 55753 are generated in above figure no 11.Fixed machine frame at end as shown in figure 12.
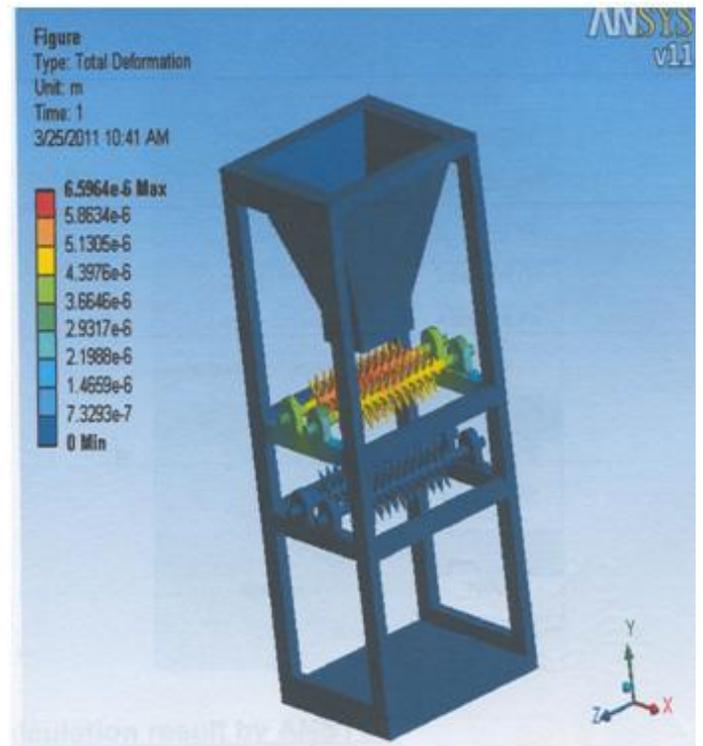
Figure 13. Total Deformation
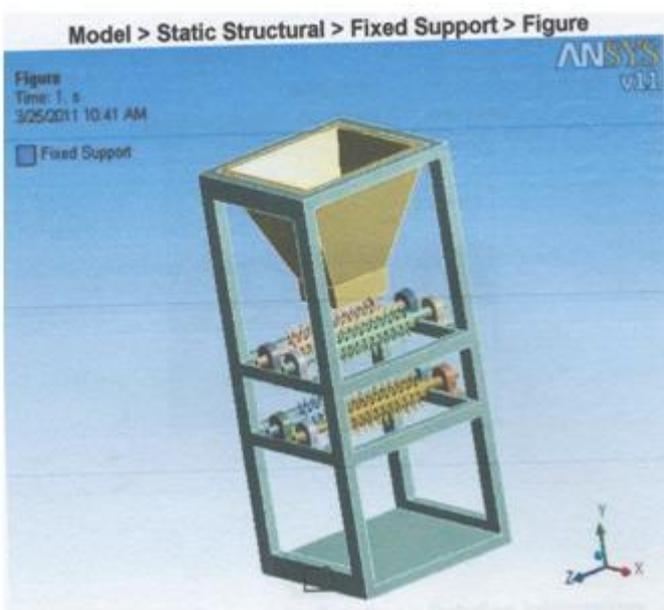
Figure 12.Boundary Condition

The finite element analysis of machine as per loading and boundary condition revealed the stress distribution in form of stress contours and total deformation as shown in figure no 13 and Figure 14.
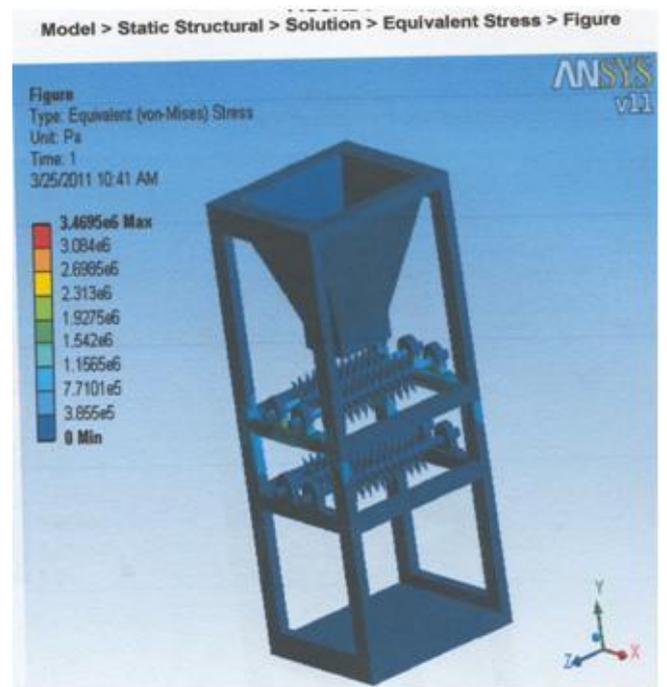
Fig.14.Equivalent stress

A. Salient Features

- Robust machine of all steel construction.
- All contact parts are made up of Stainless Steel or other food grade material.
- Speciaf1y designed S .S need1e platform to punch the amlas.
- Uniquely designed drive in order to convey the amlas in forward direction as well as to run the needle platform.
- Manual feeding of the amlas.
- Arrangement for pricking of amlas of all sizes.
- No damage of the seed.

B. Advantages

- It increases the production rate.
- It reduces the time consumption.
- It is compact in shape.
- The efficiency of machine is more.
- No skilled labour is required.
- It is cheaply for mass production.
- It can punch all the size of amlas easily.
- It requires less power for operating.
- It minimizes the problem of manual accident.

C. *Disadvantages*

- Possibility of corrosion &wearing of needle

IV.   CONCLUSION

Amla punching machine is efficient to punch the amlas which is used for prepare the murraba. This is the modern technology mentioned in the study for preparation of amla's murabba is hygienic, consume lesser time and provide maximum retention of nutrients. On the basis of design and analysis by PRO-E and ANSYS software we conclude that the stresses occurred on the machine is under control. Hence this machine is safe with in respect of operators and environment.

REFERENCES

[1] N.K.Ingole,D.V. Bhope "Stress analysis of chassis of tractor tariler for self weight redution" Ijest Vol 3 No .9 7218-7225

[2] A. Ertu¨ rka, H.N. O¨ zgu¨ vena, E. Budakb"Effect analysis of bearing and interface dynamics on tool point FRF for chatter stability in machine tools by using a new analytical model for spindle–tool assemblies"

[3] Gyu-Hong Kang, Jung-Pyo Hong, "Improved Parameter Modeling of Interior Permanent Magnet Synchronous Motor Based on Finite Element Analysis" Member, IEEE, Gyu-Tak Kim, and Jung-Woo Park ieee transactions on magnetics, vol. 36, no. 4, july 2000

[4] Min Hu a,*, Zhongqin Lin a, Xinmin Lai a, Jun Ni b, "Simulation and analysis of assembly processes considering compliant, non-ideal parts and tooling variations" International Journal of Machine Tools & Manufacture 41 (2001) 2233–2243

[5] S.S.Sane, Ghanashyam Jadhav, Anandaraj.h " Stress analysis of light commercial vehicle chassis by FEM" piaggio vehicles pvt ltd pune.

[6] Ashutosh Dubey, Vivek Dwivedi," Vehicle Chassis Analysis: Load Cases &Boundary Conditions for Stress Analysis"

[7] Tirupathi V. Chandrupatle, Ashok D.Belegundu "Introduction to Finite Elements in Engineering", 6 th edition, 2008, Pearson education, New Delhi

[8] Manual: User Manual, 'Ansys work bench 11'.

Power Quality Improvement Using Active and Passive Power Filters

C.NALINI KIRAN

(Department of EEE, G.Pullaiah College of Engineering and Technology, Kurnool, India)

ABSTRACT

Power quality standards (IEEE-519) compel to limit the total harmonic distortion within the acceptable range. Active power filter which has been used there monitors the load current constantly and continuously adapt to the changes in load harmonics. Hybrid active filter with proposed control algorithm for three phase hybrid power filter is studied here .It is composed of series active filter connected in series to the line and passive filter connected in parallel with the load . Traditionally, a passive LC power filter is used to eliminate source current harmonics when it is connected in parallel with the load and series active filter will compensate the voltages in the line. The proposed control algorithm is based on the generalized p–q theory. It can be applied to both harmonic voltage injection and harmonic current injection and also it improves the behavior of the passive filter. This control algorithm is also applied to shunt active power filter, combination of series active and shunt active and comparative study has been done. Simulations have been carried out on the MATLAB SIMULINK platform with different filters and results are presented.

Keywords **-** **Active filter, Dual-instantaneous power theory, hysteresis controller, harmonics,non-linear load.**

I.   INTRODUCTION

A power-quality problem is an occurrence manifested in a nonstandard voltage, current, or frequency deviation that results in a failure or a disoperation's of end-use equipment. Power quality is a reliability issue driven by end users. There are three concerns.   The characteristics of the utility power supply can have a detrimental effect on the performance of industrial equipment.

Harmonics produced by industrial equipment, such as rectifiers or ASDs, can have a detrimental effect on the reliability of the plant's electrical distribution system the equipment it feeds, and on the utility system.

The characteristics of the current and voltage produced by ASDs can cause motor problems. While power quality is basically voltage quality, it is not strictly a voltage issue. Since the supply system has a finite, rather than an infinite, strength, currents outside the direct control of the utility can adversely affect power quality. These are harmonic load currents, lightning currents, and fault currents.

Voltage sag is a short-term, few-cycles duration, drop in voltage on the order of more than 10% to less than 90%. Typically, it lasts from 0.5 cycles to a minute. Voltage sags result from the voltage drop, from starting big motors across-the-line, or from a fault on an adjacent power line. Voltage swell is a short-term increase in voltage of a few cycles duration. The magnitude of the increase is more than 10% and less than 80%. A swell can result from a single line-to-ground fault that raises the voltage on the other two phases. It can also result from dropping a large load or energizing a capacitor bank.

II.   NEED FOR HARMONIC COMPENSATION

The implementation of Active Filters in this modern electronic age has become an increasingly essential element to the power network. With advancements in technology since the early eighties and significant trends of power electronic devices among consumers and industry, utilities are continually pressured in providing a quality and reliable supply. Power electronic devices such as computers, printers, faxes, fluorescent lighting and most other office equipment all create harmonics. These types of devices are commonly classified collectively as 'nonlinear loads'. Nonlinear loads create harmonics by drawing current in abrupt short pulses rather than in a smooth sinusoidal manner. The major issues associated with the supply of harmonics to nonlinear loads are severe overheating and insulation damage. Increased operating temperatures of generators and transformers degrade the insulation material of its windings. If this heating were continued to the point at which the insulation

fails, a flashover may occur should it be combined with leakage current from its conductors. This would permanently damage the device and result in loss of generation causing widespread blackouts.

One solution to this foreseeable problem is to install active filters for each nonlinear load in the power system network. Although presently very uneconomical, the installation of active filters proves indispensable for solving power quality problems in distribution networks such as harmonic current compensation, reactive current compensation, voltage sag compensation, voltage flicker compensation and negative phase sequence current compensation. Ultimately, this would ensure a polluted free system with increased reliability and quality.

The objective of this project is to understand the modeling and analysis of a shunt active power filter. In doing so, the accuracy of current compensation for current harmonics found at a nonlinear load, for the PQ theory control technique is supported and also substantiates the reliability and effectiveness of this model for integration into a power system network. The model is implemented across a two bus network including generation to the application of the nonlinear load.

The aim of the system simulation is to verify the active filters effectiveness for a nonlinear load. In simulation, total harmonic distortion measurements are undertaken along with a variety of waveforms and the results are justified accordingly.

III. NEED OF SHUNT ACTIVE FILTER
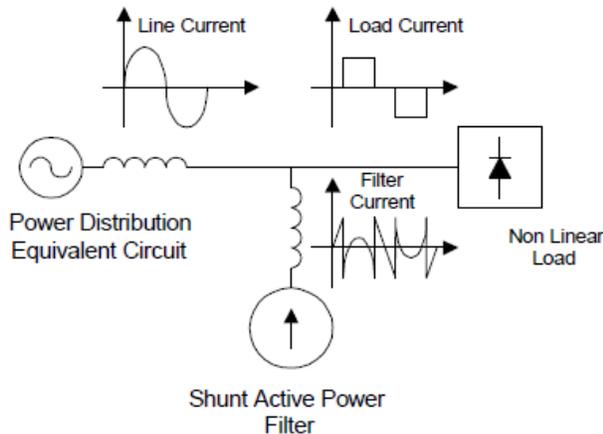
Fig 1: Need of shunt active filter

The inverter in the Shunt Active Power filter is a bilateral converter and it is controlled in the current Regulated mode i.e. the switching of the inverter is done in such a way that it delivers a current which is equal to the set value of current in the current control loop. Thus the basic principle of Shunt Active Filter is that it generates a current equal and opposite to the harmonic current drawn by the load and injects it to the point of coupling there by forcing the source current to be pure sinusoidal. This type of Shunt Active Power Filter is called the Current Injection Type APF.

Hybrid Filter

In general, the passive filter was designed only to compensate the source current harmonics; the reactive power was not considered, the concern for compensating voltage harmonics is not high due to the fact that power supplies usually have low impedance. Generally, at the point of common coupling, ridged standards are implemented to ensure a correct level of total harmonic distortion (THD) and voltage regulation is maintained. The problem of compensating for voltage harmonics is to ensure the supply to be purely sinusoidal. This is important for harmonic voltage sensitive devices such as power system protection devices and superconducting magnetic energy storage. Voltage harmonics are related to current harmonics by the impedance of the line. Although compensation of voltage harmonics helps to provide a reduction in current harmonics, this however, does not negate the necessity to current harmonic compensation.
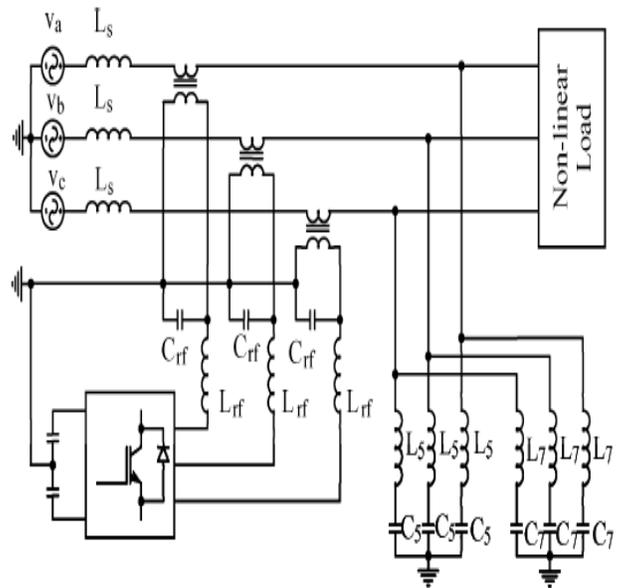
Fig 2: Block diagram of hybrid filter

IV. DUAL INSTANTANEOUS POWER THEORY

The proposed control algorithm is based on the generalized p–q theory. It may be applied to both harmonic voltage injection and harmonic current injection. In this algorithm, the compensation voltage references are extracted directly. Therefore, the calculation of the compensation voltage reference will be much simpler than for other control algorithms. In addition, the difficulty of finding the voltage reference gain disappears.
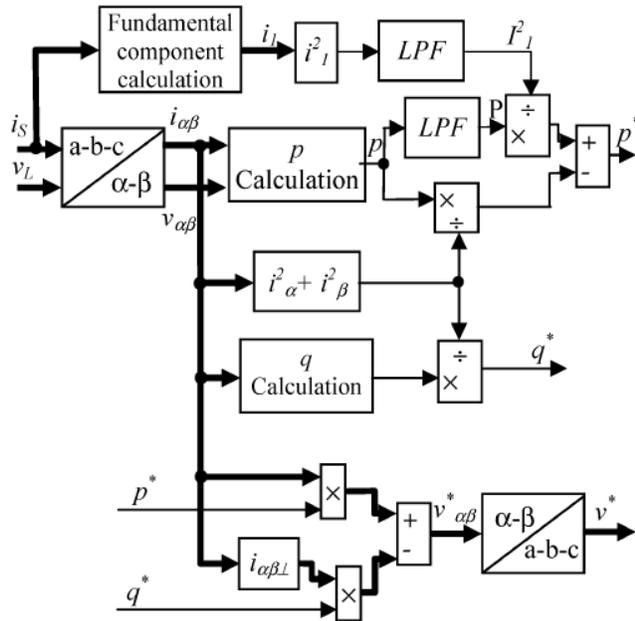
Fig 3: Dual instantaneous power theory
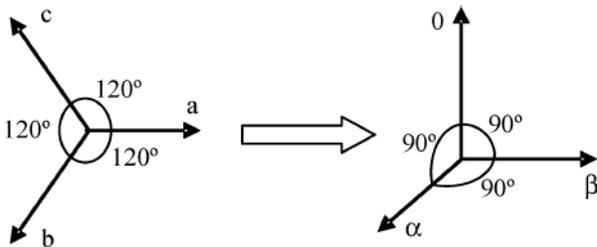
Fig 4:Three phase to two phase conversion

V. RESULTS AND DISCUSSIONS

Here simulation has been carried out for series active, shunt active, shunt passive, filters, by using MATLAB SIMULINK..FFT analysis is done, THD is observed for various circuits.
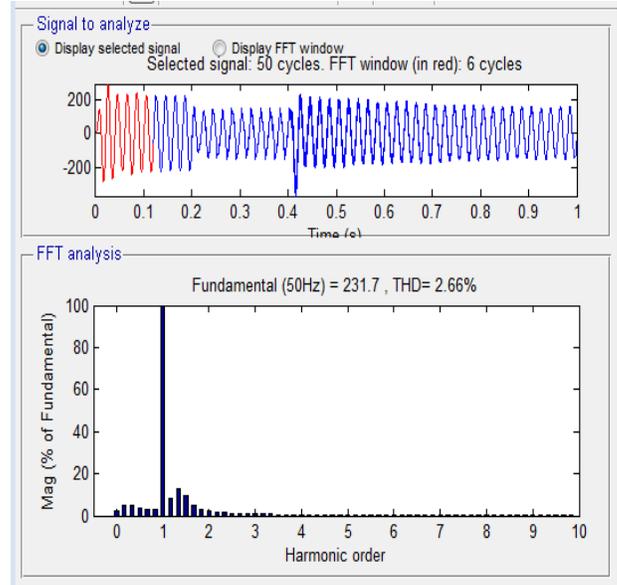
Fig 5: FTT analysis of compensating voltage wave form
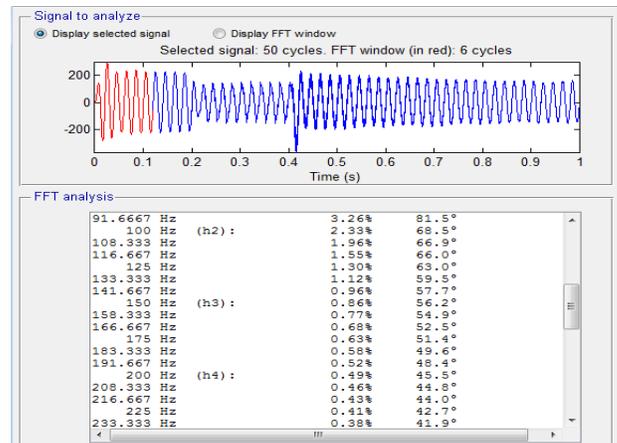
Fig 6: Representation of each and every harmonic
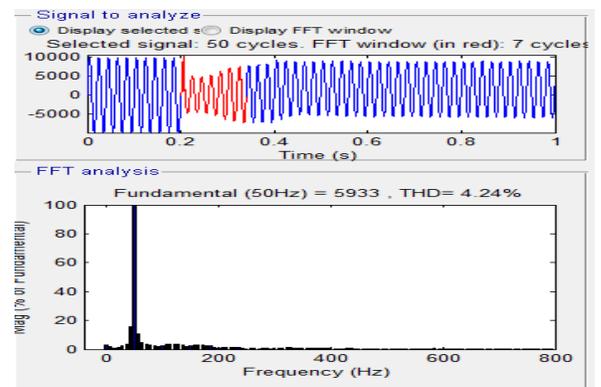
Fig 7: FFT analysis of compensating current
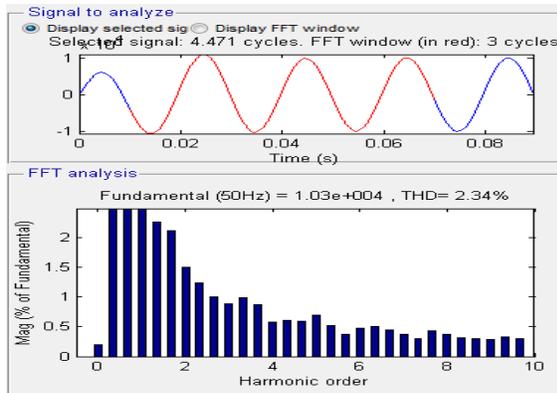
Fig 8: FFT analysis of Compensating voltage wave form at PCC

VI. CONCLUSION

A dual instantaneous power theory based on instantaneous power theory for hybrid power filters is studied, a Simulink model is designed and total harmonic distortion is calculated using FFT analysis.hybrid power filter which has been used here monitors the load current constantly and continuously adapt to the changes in load harmonics. The performance of three phase hybrid power filter using dual instantaneous power theory is simulated. This control algorithm is also applied to shunt active power filter, combination of series active and shunt active and comparative study has been done. Simulations have been carried out on the MATLAB SIMULINK platform with different filters and results are presented.

REFERENCES

[1] F. Z. Peng and D. J. Adams, "Harmonics sources and filtering approaches,"in *Proc. Industry Applications Conf.*, Oct. 1999, vol. 1, pp.448–455.

[2] H. L. Ginn, III and L. S. Czarnecki, "An optimization based method for selection of resonant harmonic filter branch parameters," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1445–1451, Jul. 2006.

[3] B. Singh, K. Al-Haddad, and A. Chandra, "A review of active filters for power quality improvement," *IEEE Trans. Ind. Electron.*, vol. 46, no. 5, pp. 960–971, Oct. 1999.

[4] S. George and V. Agarwal, "Optimum control of selective and total harmonic distortion in current and voltage under non sinusoidal conditions," *IEEE Trans. Power Del.*, vol. 23, no. 2, pp. 937–944, Apr. 2008.

[5] H. Akagi, Y. Kanazawa, and A. Nabae, "Generalized theory of the instantaneous reactive power in three-phase circuits," in Proc. IEEE and JIEE IPEC, 1983, pp. 821–827.

[6] Generalized theory of the instantaneous reactive power in three-phase circuits," in Proc. IEEE and JIEE IPEC, 1983, pp. 821–827.

[7] A. Luo, Z. Shuai, W. Zhu, R. Fan, and C. Tu, "Development of hybridactive power filter based on the adaptive fuzzy dividing frequency-control method," *IEEE Trans Power Del.*, vol. 24, no. 1, pp. 424–432, Jan.2009.

[8] R. S. Herrera and P. Salmerón, "Instantaneous reactive power theory: A comparative evaluation of different formulations," *IEEE Trans. PowerDel.*, vol. 22, no. 1, pp. 595–604, Jan. 2007.

[9] Schlabbach, D. Blume, and T. Stephanblome, Voltage Quality in Electrical Power Systems, ser. PEE Series. New York: IEE Press,2001.

[10] L. Gyugyi and E. C. Strycula, "Active AC power filter," in Proc. IEEE IAS Annu. Meeting, 1976, pp. 529–529.

[11] Y. Komatsu and T. Kawabata, "Experimental comparison of pq and extended pq method for active power filter," in Proc. EPE, 1997, pp.2.729–2.734

[12] V. Soares, P. Verdelho, and P. D. Marques, "Active power filter control circuit based on instantaneous active and reactive current id – iq method," in Proc. IEEE PESC, 1997, pp. 106–101.

[13] Generalized theory of the instantaneous reactive power in three-phase circuits," in Proc. IEEE and JIEE IPEC, 1983, pp. 821–827.

[14] J. W. Dixon, G. Venegas, and L. A. Moran, "A series active power filter based on a sinusoidal current-controlled voltage-source inverter," *IEEE Trans. Ind. Electron.*, vol. 44, no. 5, pp. 612–620, Oct. 1997.

[15] H. Yang and S. Ren, "A practical series-shunt hybrid active power filter based on fundamental magnetic potential self-balance," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2089–2096, Oct. 2008

[16] .H. Akagi, "Active harmonic filters," *Proceedings of theIEEE* Volume 93, Issue 12, Dec. 2005 Page(s):2128 –2141

[17] J.C. Das, "Passive Filters-Potentialities and Limitations", *IEEE Trans. on Industry Applications*, Vol. 40, no. 1, January 2004, pp. 232-241.

Co-Operation between Librarians and PhD Students to Enhance Library Collection Development

Dr. Qais Faryadi
Faculty of Science and Technology, Department of Computer Science
Universiti Sains Islam Malaysia (USIM)

ABSTRACT
This study investigates whether there is adequate communication between librarians and the postgraduate students in the collection development process. The primary objective of this research is to critically evaluate the existing collection development process of *Pandai* librarians and investigate the criteria used to strengthen the library collection development. This research also critically investigates whether there is effective communication between librarians and postgraduate students to improve the library collection development. A triangulation method (quantitative, qualitative and descriptive) is employed in the investigation. The instruments used in this study, in which 150 postgraduate (PhD) students, 14 librarians and 15 support staff of the library participated, included questionnaires, interviews and observations. Result from observation revealed that librarians were helpful in providing information at the reference counter. Their body language, smiles and eye contacts communicated friendliness. Overall observations showed that reference librarians and support staff were well mannered, sociable and responsive. The findings from interviews and questionnaires showed that there was inadequate communication between the librarians and postgraduate students (PhD) to enhance the collection development process. The majority of the support librarians 83.3% agreed that postgraduate students were not consulted for the collection development process while only (16.6 %) indicated that postgraduate students were sometimes consulted by email. From 12 qualified librarians with a master's or basic degree, 85.5% of the qualified librarians said that in their collection development process, postgraduate students (PhD) were not consulted.
Index Terms: Postgraduate students, library, collection development, communication, decision making

Introduction
Communication between librarians and library users, especially postgraduate students, is vital in enhancing the library's collection development process as well as meeting the scholarly needs of students. A very crucial function of a library is its collection development, which also impacts the status of the university, nationally and internationally (Zain,2004). Effective communication between librarians and postgraduate students allows librarians to evaluate their own effectiveness. Librarians and knowledge seekers alike are responding to this core issue to enhance their collection development (Kathleen et al., 2007). Information and communication technologies have paved the way for librarians and users to communicate with each other in order to provide vital feedback and much needed information concerning their work. Internet and e-mail are very useful tools for effective communication for such a purpose. Librarians and the students must be educated to cooperate in order to achieve successful information delivery (Gudakuvaska, 2001). As such, scholarly communications are forcing librarians to shift their mental models and modify their services accordingly. Librarians have to attune their vision and mission to the needs of library users. Consultations with users would certainly help to fine-tune achievable goals and objectives. Only then can librarians celebrate success in their scholarly collections and development.

Problem Statement
This study investigates whether there is adequate communication between librarians and the postgraduate students in the collection development process. It is important that librarians attract users by increasing their value added activities in the library (Taylor, 1986). Good communication would also pave the way to meeting the scholarly needs of students (Normann, 1994). Librarians must communicate with students to improve their collection while students must come forward to discuss their needs. Such collaborative discussions or joint activities would create a meaningful learning environment (Suchman,1987; Davenport, 2002). To ensure that librarians improve their collection to cater to the needs of its users, especially PhD students, it is important to explore the factors that affect the quality of their collection. In this connection, it would be useful to know how the librarians strengthen their collection and development of materials.

Literature Review
For librarians to stand tall in their collection development, they must consider their users' needs as well as be in constant communication with library users, especially postgraduate students (Durr, 2011). Postgraduate students require more scholarly materials to do better research. Librarians need to consult them regularly and update their material collection.

Only then, librarians can offer engaging and purposeful materials to their patrons. Librarians are, undoubtedly, a vital part of a systematic delivery system. Librarians are not merely book keepers; more importantly, they should function as information managers. They need to convey information not only within the library but beyond the library walls too (Berry, 2011) High quality customer service must be part and parcel of their responsibilities (Paterson 2011). Apart from being custodians of information, librarians must be ready to provide users, especially postgraduate students, with information that would help them in their scholarly research. One of the major functions of a good library is to provide life-long learning. Hence it is essential that librarians and library users, especially postgraduate students, communicate on the same wavelength (Iveta, 2001).

Librarians should always strive to improve their efficiency in providing useful information to meet the needs of their clientele. There should be opportunities for interaction between librarians and users. Librarians must ensure all communication devices are utilized in the best manner possible for information acquisition, presentation and communication (ESF, 2002). As an information provider, one of the top objectives of a library must be to ensure that its users are able to gain maximum benefit from the resources available. This objective can be achieved by investing in efforts to educate users and library staff alike. Staff education includes improving their qualifications, providing conducive and motivating work environment and, most importantly, setting higher performance targets. Competent librarians also stay connected with their clientele. When both parties are able to communicate effectively, scholarly research and other academic pursuits will be more successful.

Communication between librarians and students, especially postgraduate students, is vital if the library wants to ensure that it achieve its goals (Mary, 2010). Librarians play a very significant role in our society. They help users find much needed information for their research. Apart from assisting end-users, the responsibilities of librarians include document processing, sourcing information, and providing strategic planning and training. These vital functions of librarians can only be carried out successfully if there is adequate communication between students and librarians regarding the library needs of the former (Doug, 2010).

Significance of the Study
1. The findings of this research would help librarians improve communication with users, especially postgraduate students, in order to enhance the library's collection development process.

2. The findings of this research would help librarians make better investment regarding the purchase of communication devices.
3. The findings of this research would also help the university staff and scholars locate the required reference materials more efficiently.
4. Feedback from observations, interviews and questionnaires provides valuable information for future researchers.

Objectives
1. To critically evaluate the existing collection development process of librarians and investigate the criteria used to strengthen the library collection development.
2. To critically investigate whether there is effective communication between librarians and postgraduate students to improve the library collection development.
3. To propose a viable and workable plan of action to librarians so that resources are used efficiently.

Methodology and Data Collection
This study applied quantitative and qualitative analyses coupled with observations to carry out the investigation. Since this research deals with human feelings and perceptions, such a triangulation method is most appropriate in conducting the study. Research also shows that a combined methodology design helps the researcher to better understand the research problem (Faryadi, 2010). The data collection for this study was conducted by the researcher. The whole process of distributing the questionnaires on and collection of feedback took two weeks.

The response rate was satisfactory. From 150 questionnaires distributed to postgraduate students, 120 students responded. 15 librarians were requested for an interview; 12 attended. 15 support librarians from the reference desk agreed to attend an interview. The interviews were conducted informally to elicit better cooperation. The questionnaires and interview sessions had been tested for reliability and validity on a similar population i.e. pilot group. The questioners were then fine-tuned and retested with the target group.

The framework for data analyses in this study was adapted from the work developed by Miles and Huberman (1994). The primary tools used to analyze the collected data were the Statistical Package for Social Sciences (SPSS version 16.0) and Microsoft Excel. Data were categorized and meaningfully reconstructed according to the problem statements and objectives of the research. Crucial data were selected for scrutiny, before being simplified for easy comprehension.

Data were then cross-checked few times to determine its face validity and reliability.

Population and Instruments

This research was conducted at the Library of *Pandai* in Malaysia. At the point of writing, there were more than (150) postgraduates pursuing PhD programs. A total of 150 postgraduate students (PhD) and 14 librarians participated in the study. Apart from the 14 librarians, the chief librarian, deputy chief librarian and 15 support staff agreed to be interviewed. The instruments used to collect data were (1) questionnaires for students; (2) questionnaires for librarians; (3) interviews with randomly selected students and librarians; (4) interviews with the chief librarian and his deputy; (5) observations of activities at the reference desk in the library; and (6) observations of the librarians' daily activities in the library.

Results

Result from observation revealed that librarians were helpful in providing information at the reference counter. Their body language, smiles and eye contacts communicated friendliness. Overall observations showed that reference librarians and support staff were well mannered, sociable and responsive. The findings from interviews showed that there was inadequate communication between the librarians and postgraduate students (PhD) to enhance the collection development process. The main line of communication was through supervisors and lecturers who recommended reference materials for students, especially those pursuing postgraduate studies.

This study further illustrated that postgraduate students (PhD) were not consulted regarding their scholarly needs. The majority of the support librarians (83.3%) agreed that postgraduate students were not consulted for the collection development process while only16.6 % indicated that postgraduate students were sometimes consulted by email. From 12 qualified librarians with a master's or basic degree, 85.5% of the qualified librarians said that in their collection development process, postgraduate students (PhD) were not consulted. Observations and surprise checks at the reference desk indicated that books requested were not always available.

During the two weeks of observation at the reference desk, it was noted that activities in the reference desk were fewer than usual. May be it was due to the revision week for the exam. Feedback from the postgraduate participants (PhD) showed that majority of them (80%) indicated that they had not been invited to participate in any discussion on collection of library materials.

It was noted that only 4.1% of the postgraduate (PhD) students did not answer this question.

Table 1 Summary of the Research Instruments and the Respondents

No	Instruments	Respondents	Sample	Collection
1	Questionnaire	Students	150	120 (80%)
2	Interviews	Librarians	14	12 (85.7%)
3	Observation	Library- students	300	300 (100%)
4	Interview	Reference desk	18	10 (83.3%)

Summary of the Research Instruments and the Respondents

Table 1 shows the summary of data collection instruments and the participants. The results indicated that out of 150 PhD students 120 (80%) agreed to participate in the research. Out of 14 librarians only, 12 participated while out of 18 support librarians, 10 agreed to be interviewed. Observation from communication between support librarians and students indicated that reference librarians were pleasantly helpful in providing needed information to the students. Unstructured interviews were conducted to determine librarians and support staff performance.

It is worth mentioning that the questions tendered to the participants were not formal questionnaires. The purpose of having informal questions was to encourage the interviewees to give as much information as possible.

Table 2: Summary of the Collection Development Process

Selection of material	Acquisition	Assessment	Weeding	Budget
Academic staff Supervisors Dean's approval Library staff Community interest Popularity of the item Subscription	Best possible price Catalogue and Media jobbers Local venders Standing orders	Needs assessment Continual evaluation Person responsible	Periodically removed Reason for weeding Criteria for removing Donate to schools	Depend on the Ministry of Finance approval

Summary Of collection development process

Table 2 explains the process whereby USIM library strengthens its collection and development of materials. The most important aims and objectives of this research were to find out how *Pandai* library processes its collection development. This research also investigates whether *Pandai* library communicates with PhD students in order to meet their scholarly needs and to strengthen their library collections.

Interviews conducted with the reference librarians indicated that most of the collection development process carried out based on recommendations from the academic staff, supervisors and the deans of the faculties

for final approval. Unfortunately, this research found that there was no consultation and communications between the librarians and the PhD students to reinforce the library collection development. Interviews also indicated that the library's collection development process was also based on community interests, popularity of the materials and subscriptions.

The reference librarians were asked this question: H*ow do the librarians acquire materials*? They replied that their acquisitions were mostly based on the best price possible method, vendors' catalogues, media jobbers, local and international vendors and standing orders. This research further showed that the process of assessment of materials was done by a library staff whose job was to continually evaluate the library materials for weeding, after making sure that there was a copy for future reference. The staff was required to give the reasons in writing for his/her decision to remove a particular book. It is worth mentioning that sometimes instead of disposing of the materials; the books were donated to public or school libraries. In response to the question: *How do you manage your budget for the purchase of materials,* the collection development librarian replied that the Ministry of Finance had the final say in approving or disapproving the budget.

Table 3: Reference and Circulation Desks Observation

No	Evaluation Criteria	Responses
1.	Was there anyone present at the reference desk?	Yes
2.	How many people approached the reference desk per hour?	15 per hour
3.	Was the reference librarian interested to help?	Yes
4.	Did the reference librarian look bored or enthusiastic?	Enthusiastic
5.	Did the reference librarian appear disturbed?	No
6.	Did you find what you asked for?	No
7.	How much time did the reference librarian spend with a user?	less than 5 min
8.	Did the reference librarian clarify user's questions?	No
9.	Did the reference librarian make value judgment regarding users' inquiry?	No
10.	How did the reference librarian respond to queries: print or electronic, or both?	Electronic

Reference Desk Observation Results

Table 3 indicates the observation checklists to evaluate the activities at the reference desk. Observation revealed that librarians were helpful in providing information at the reference counter.   The librarians were quite enthusiastic in their job and were more than ready to render help to users. The Library was arranged in a way that books, reference desk, computers and self-help machines are seen easily by the users. The reference desk is located in the middle of reference materials and users sitting area so that it is easily noticed from far. A total of 8 self-check machines and 16 computer terminals have been installed to assist the users meet their information needs. The observation signified that two support librarians operated the circulation counter while a librarian was on duty at the reference desk during the library opening hours.

Observation further showed that the librarians at the reference desk and support librarians at the circulation desk patiently waited for patrons to approach them. Observations revealed that librarians were aware of the approaching patrons and immediately offered their assistance. Their body language, smiles and eye contacts communicated friendliness. Each patron was greeted courteously with *Selamat pagi, boleh saya bantu? (Good morning, may I help you?)* Overall observations showed that reference librarians and support staff were well mannered, sociable and responsive.

Table 4: Summary of Important Questions for Librarians

	Do you communicate with postgraduate students to strengthen your collection development?	Do you think Communication with postgraduate students is vital for survival of the library?	Do you educate the users on how to use the library?
Yes	(0%)	4 (33.3%)	7 (58.3%)
No	12 (100%)	8 (66.6%)	
Sometimes	0%		
Once or twice	0%		5 (41.6%)

Summary of Important Questionnaires for Librarians

Table 4 shows the majority of the librarians (85.7%) interviewed indicated that they did not consult postgraduate students (PhD) in their collection development process. Only lecturers and supervisors were invited to advise them. However, regarding a question whether communication between them and postgraduate students was vital, the majority (66.6%) were of the view that it was not vital.

Only 33.3% of the respondent agreed that communication was vital. When asked whether they had taught students how to use the library, 58.3% said they had frequently done so, while 41.6% replied that they had educated users once or twice. It is interesting to note that the majority of the librarians and support librarians were of the view that experience was more important than qualifications in the running of the library.

Table 5: Summary of Important Questions for Students

	Have you ever been consulted by the librarians to participate in collection development process?	Do you think communication between librarians and users is vital?	Have you been trained by the librarians how to use library materials?
Yes	0%	120 (100%)	20 (16.6%)
No	120 (100%)	0%	85 (70.8%)
Sometimes	0%	0%	5 (4.1%)
Once or twice	0%	0%	10 (8.3%)

Summary of Important Questionnaires for Students

Table 5 presents the summary of important questions for students. The results showed that 80% of the respondents indicated that they had not been consulted in any collection development process. To a question whether communication between librarians and students was vital, the majority of the students (80%) said it was. The results further indicated that 13.3% of the respondents said that they had received training on how to use library materials while 56.6% said they had not. It is of interest to note that 6.6% of the respondents indicated that they been trained at least once or twice.

Conclusion and Discussion

This study obtained valuable feedback regarding the collection development process at the *Pandai* library. Observation revealed that librarians were helpful in providing information at the reference counter. Their body language, smiles and eye contacts communicated friendliness. Overall observations showed that reference librarians and support staff were well mannered, sociable and responsive. However, the research revealed that there was no cooperation and consultation between postgraduate (PhD) students and the library collection development department. It is unfortunate that PhD students were not consulted in the collection department process. The success of collection development of librarians depends on whether it is able to meet the scholarly needs of users. The collection development process in *Pandai* library should not consult only lecturers and supervisors. Feedback from PhD students, who have to carry out extensive research, is vital. Perhaps *Pandai* librarians need to compare notes with librarians in other more established libraries to improve the material collection process.

Implications of the Study for *Pandai* Library Policy Makers

Pandai library policy makers have an important role to play in improving the collection process. Students, lecturers as well as the Education Ministry must support *Pandai* library policy makers to redesign the concepts of collection development process. The findings of this study showed that the lack of communication between librarians and PhD students did not augur well for the development of the collection process. In addition, policy makers should support *Pandai* library planners by providing them with enough funds to continue their collection development effectively.

The following are some suggestions for *Pandai* library policy makers to re-evaluate their collection development:
1.   Support and create a library environment that focuses on meeting the scholarly needs of post-graduate students.
2.   Support the creation of a library environment that invites scholarly engagement.
3.   Support the allocation of funds for scholarly materials.
4.   Support autonomy by the collection development department in decision making regarding purchases of materials.

Implications for Future Research

This research investigated the relationship between postgraduate (PhD) students and the library collection development process in *Pandai* library. However, it would be useful to obtain feedback from other library users, besides PhD students, regarding the library collection development. Further research is necessary to investigate *Pandai* library collection development process and compare it with other libraries to find out how they process their collection development so that *Pandai* library may benefit from them.

Limitations of the Study

The results and conclusions of this study were subjected to the following limitations: First, this research focused only on postgraduate (PhD) students. Thus, the results cannot be generalized and further investigations are needed to examine how other users such as Masters and undergraduate students can bring a positive change in the library collection development process. Second, this research recommends further investigations involving longer periods of observation as well as a larger sample to evaluate the effectiveness of material collection. Further, as some of the data were obtained through observations, the subjective bias of the researcher was inevitable.

Suggestions for Further Improvement of the Library

1.   Create a learnable, user-friendly and effective library web site to provide direction to the users.
2.   Encourage users to submit feedback and suggestions online.
3.   Prototype "See You See a Librarian" developed in 1996 by Eric Lease Morgan. The purpose of this service was to investigate the possibility of

providing chat communication between librarians and between librarians and patrons.

4. In the library students are the primary users of digital reference, and they tend to prefer chat reference service to e-mail because it involves a two-way conversation in real time, very much like talking to a reference librarian in person. Chat users can receive immediate feedback, thus they can use written language in the same manner used in a person-to-person conversation.

5. There are several chat systems librarians can choose from; the most widely known are 24/7 Reference, LSSI – Virtual Reference Toolkit, and Question Point. Although such services are commonly associated with a 24-hour-a-day, seven-days-a-week service, they can be modified to offer reference service at specific times.

6. With chat, librarians can use a variety of tools to facilitate communication with the patron. One of the most important features in chat reference is the use of software with the ability to co-browse. This feature allows the librarian and the user to communicate while viewing the same web pages.

7. To respond quickly, the librarian can also use pre-written messages. These messages involve typical greetings and sign off texts and are used to reduce the time and typing involved in the reference interview.

References

[1] Berry, John N., III, (20011). Nancy Pearl: LJ's 2011 Librarian of the Year, Library Journal, v136 n1 p24-26.

[2] Davenport, E. & Hall, H. (2002) Organizational knowledge and communities of practice. Annual Review of Information Science and Technology, 36, 171-227.

[3] Doug, (2010). From Librarian to Knowledge Manager and Beyond: The Shift to an End-User Domain, Phase   5 Consulting   Group Inc.

[4] Durr, C. (20011). Making Wise Buys: Five Values to Consider when Evaluating a Library Purchase, Computers in Libraries, v31 n6   p6-10

[5] ESF (2000-2002) | Course Modules the Library Association and Canterbury Christ   Church University College. ICT   INSET   for Librarians

[6] Faryadi, Q. (2010). Developing an Effective Interactive Multimedia Instructional Design to Teach   Arabic Language, p. 80, Al-Mehrab e-Publisher, Kuala Lumpur, Malaysia,

[7] Gudakuvaska, I, (2001). Focus on further education of librarians in Latvia.   Information Research, Vol. 7, No. 1

[8] Iveta G. (2001) Focus on further education of librarians in Latvia. Information   Research, Vol. 7 No. 1,

[9] Kathleen A. Newman, Deborah D. B.ic, & Kimberly L. Armstrong, (2007). Scholarly Communication Education   Initiatives SPEC Kit 299 (Washington,   D.C.: Association of Research Libraries,

[10] Mary, W., Jordan (2010). All Stressed Out? Enumerating and Eliminating Stress in the Academic Library, Brick and Click   Libraries Symposium Proceedings.

[11] Miles, M.B, and Huberman, A.M. (1994). Qualitative Data Analysis, 2nd Ed., p.   10-12. Newbury Park, CA: Sage.

[12] Normann, R. & Ramirez, R. (1994). Designing interactive strategy: from value chain to value constellation. Chichester, UK: John   Wiley.

[13] Paterson, N. (2011). An Investigation into Customer Service Policies and Practices within the Scottish College Library Sector: A Comparison between the Customer Service Exemplars from the Retail Sector with Current Scottish College Library Practice, Journal of Librarianship and Information Science, v43 n1 p14-21

[14] Peters, M., Roybal, S. (2011) Faculty-Library Collaboration: Embedding Information Literacy in Educational Research Graduate Classes

[15] Suchman, L.A. (1987). Plans and situated actions: the problem of human-machine communication. Cambridge: Cambridge   University Press.

[16] Taylor, R. S. (1986). Value added processes in information systems. Norwood, NJ: Ablex.

[17] Zain, M. Hawa, D. (2004). Faculty awareness on the collection development of the international Islamic University,   Vol.9, No, 2, 17-34

Note: This research was carried out in a University Library named *Pandai* Library. *Pandai* is not the real name. This research is ongoing and another five libraries will be investigated to see what their criteria are in strengthening collection development process.

STATIC AND DYNAMIC ANALYSIS ON TATRA CHASSIS

Sairam Kotari [1], V.Gopinath[2]
PG student, Associate Professor, Department of Mechanical Engineering
QIS College Of Engineering &Technology
Ongole,Andhra Pradesh

ABSTRACT

This paper deals with the analysis of chassis frame for improving its payload by adding stiffner and c channel at maximum stress region of chassis frame. The FEM analysis has been carried out with various alternatives. The results illuminate the new creative ways for optimum frame design which makes it more sustainable for structural concerns. This paper analyzed the backbone frame for both dynamic and static load condition with the stress deflection bending moment on the tatra chassis frame. The finite element analysis over ansys is performed by considering the load cases and boundary conditions for the stress analysis of the chassis. The tatra chassis is being modeled in catia v5 and then it is being imported in the finite element analysis software-Ansys. At present the payload of the tatra is 10.4 tones in this project we enhance the capacity of vehicle to 14 tones from existing chassis as per the requirement. This has been carried out with limited modifications by adding stiffeners and c channel. The necessary design changes required to enhance the load carrying capacity of the vehicle has been recommended successfully.

Keywords: Static Analysis, Dynamic Analysis, Finite Element Analysis, ANSYS, Tetra Chasis

1.0 INTRODUCTION

TATRA is mainly used in terrain conditions fitted with EURO II Engine. It can able to operate under extremely high and cold ambient temperatures, high humidity and industry environments. The Chassis of this vehicle is very rigid against torsion and bending. The chassis has high resistance to shocks and vibrations. Therefore, it protects super structures from torsion, stresses, and allows driving fast on rough roads.

This vehicle is a Left hand drive. The specialty of this vehicle is 8×8 drives, all of the eight are lockable differentials. It contains independent suspension with swinging semi-axles for every eight wheels. The leaf spring and telescopic shock absorber supports the front axle. The leaf spring alone supports the rear axle. The tyres have manually controlled central tire inflation system (CTIS) operable on the move. It is a Longer Wheel Base special Chassis. This Chassis has four axles and all the axles are of driven type. Separate axle differential provided for each axle**.**

The TATRA backbone frame consists of tubes bolted together with axle final drive housing and cross members. The backbone frame is connected through the cross member with welded steel frame. The backbone frame also protects driveline shaft from transmission to the wheels and differentials that are placed inside, against duct, moisture and outer mechanical damages

1.1 Frame

TATRA backbone frame consists of tubes bolted together with axle final drive housings and cross members. The backbone frame is connected through the cross members with welded steel frame. This solid structure is exceptionally rigid against torsion and bending.

The backbone frame also protects driveline shafts from transmission to the wheels and differentials that are placed inside, against dust, moisture and outer mechanical damages. Chassis structure gives extremely high torsion resistance protecting super structures from torsion, stresses and together with independent wheel, suspension improves mobility in rough terrain.

1.2 Engine
Engine Type   - 8 Cylinder, Turbocharged, Air cooled engine.
Engine output   - 300 KW at 1800 r.p.m.
No. of Speed   - 10 speed manual synchromesh transmission.
Engine Capacity   - 12667 cc
Maximum torque   - 1830N-m at 1200 r.p.m

1.3 Chassis

The chassis of an automobile consists of following components suitably mounted:
- Engine and the radiator.
- Transmission system, consisting of the clutch, gear box, propeller shaft and the rear axle.
- Suspension system.
- Road wheels.
- Steering system.
- Brakes.
- Fuel tank.

All the components listed above are mounted in either of the two ways, viz., the conventional construction, in which a separate frame is used and the frameless or unitary construction in which no separate frame is employed. Out of these, the conventional type of construction is being used presently only for heavy vehicles whereas for car the same has been replaced by the frameless type accept of course by small manufacturers, who still find it economical to use frame.

2.0 PHASES OF DESIGN
The complete design process, from start to finish, is highly iterative in nature and is as outlined in the flow diagram shown below. The process begins with recognition of a need and a decision to do something about it. After much iteration, the process ends with the presentation of the plans for satisfying the need. The design of products and services is accomplished in this traditional way. Fig. 2b shows a flow diagram of iterative phases.

Recognition of need

Definition of problem

Synthesis

Analysis and optimization

Evaluation

Presentation
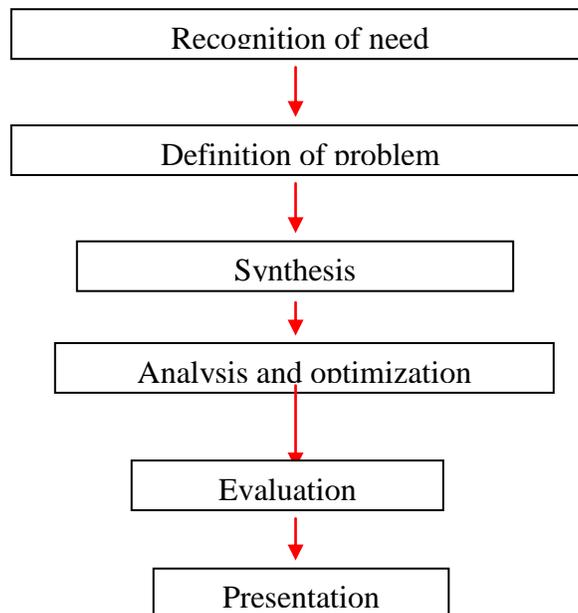
Fig. 1. Phases of design

3.0 FINITE ELEMENT ANALYSIS
The finite element analysis is a numerical analysis technique to obtain the solution of partial differential equations. The mathematical procedures such as Galerkin's weighted residual method and Raleigh-Ritz methods are used to obtain the finite element formulation of the partial differential equation. The geometrical domain describing the engineering field problem is divided into sub domains, referred to as finite elements, and the variation of the primary variable in the finite element is described using piece-wise continuous functions within each element.

4.0 GEOMETRIC MODELING OF CHASSIS

The Chassis Model of TATRA is created using CATIA V5 R16.
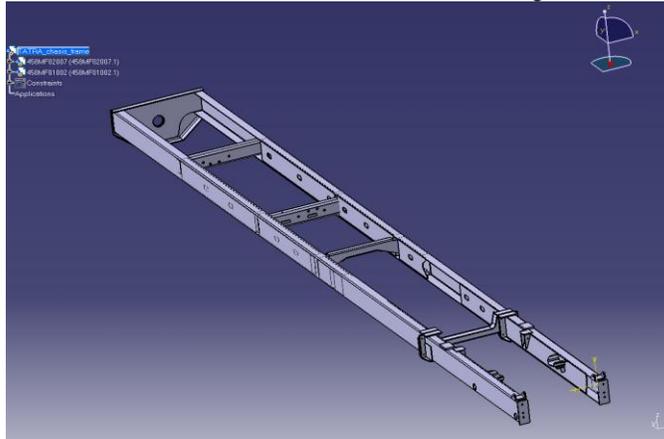
Fig.2

4.1 Material properties of chassis

No.	Material	Yield Strength (σ_y)	Ultimate Tensile Strength (σ_u)	Young's Modulus (E)	Poisson's Ratio (v)
1	High strength Structural Steel	410 N/mm^2	540 N/mm^2	2, 00,000 N/mm^2	0.3

Table 1

4.2 Specifications of the chassis

No	Description	Dimension (mm)
1	Length of Chassis	10208
2	Width of Chassis	1000

Table 2
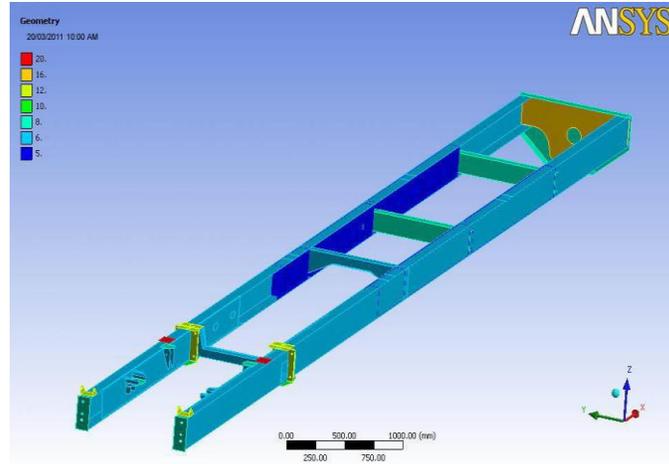
Fig 3: Ansys model with various thickness

The thickness is entered for all individual parts. It is varying from 5mm to 20mm.
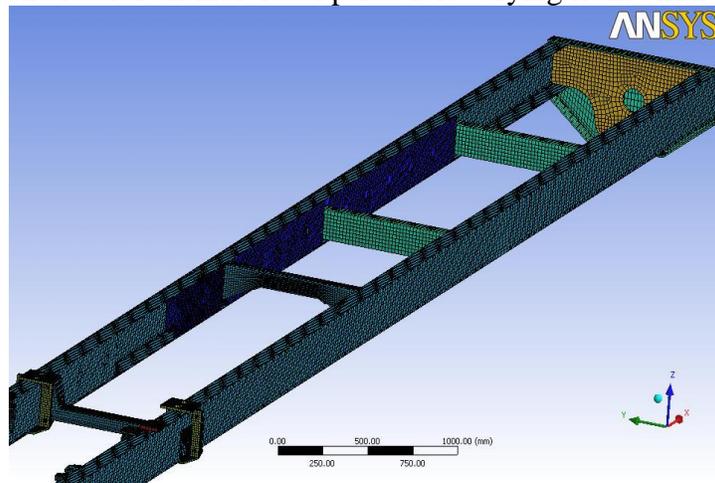
Fig 4: Finite element model of rear chassis

4.3 Boundary conditions on the chassis

S. No	Loads	Description
1.	Electronic component and Antenna Load	35000 N
2.	Cabin load	1949.7 N
3.	Engine load	3789.3 N
4.	Dynamic load	2g load applied in Z direction in centre of gravity

Table 3

5.0 ANALYSIS OF TATRA CHASSIS

> Static analysis
> Dynamic analysis
> Static analysis of the chassis:

A static analysis calculates the effects of *steady* loading conditions on a structure, while ignoring inertia and damping effects, such as those caused by time-varying loads. Static analysis determines the displacements, stresses, strains, and forces in structures or components caused by loads that do not induce significant inertia and damping effects.

6.0 RESULTS AND DISCUSSIONS
Before Modification
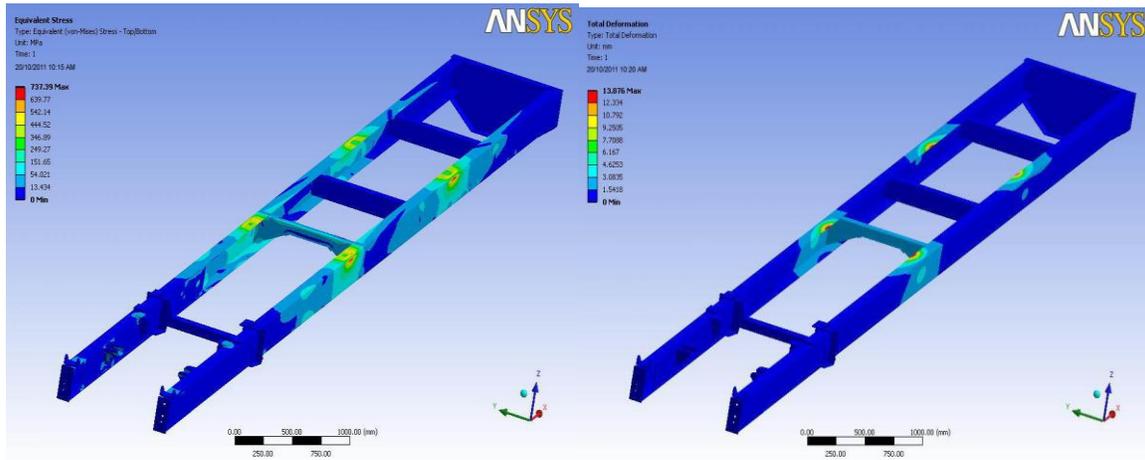
Fig 5:Stress in the Chassis   Fig 6 : Total deformation in the chassis

Max. Stress = 737.3 Mpa. / Location = just under the loading point
The above shown result plot represents the stress in the chassis.
Max. Deflection = 13.8 mm / Location = just under the loading point
The above shown result plot represents the total deformation in the chassis

From the above stress and deformation contour, stress induced in the frame is 737.3 Mpa and deformation is 13.8mm. It is more than the yield strength of the material. So it is necessary to increase the strength of the chassis frame incorporating suitable design changes.

After Modification

The strength of the chassis was increased to the safety level by adding stiffeners. Six no of stiffeners was introduced in the maximum stress induced areas which is coming in the center of the rear chassis frame. The various result plots for different thickness are shown below.

For Stiffener thickness of 4 mm:4mm thickness stiffeners was introduced in the chassis frame and static analysis was carried out and the stress and deflection contours are shown below,
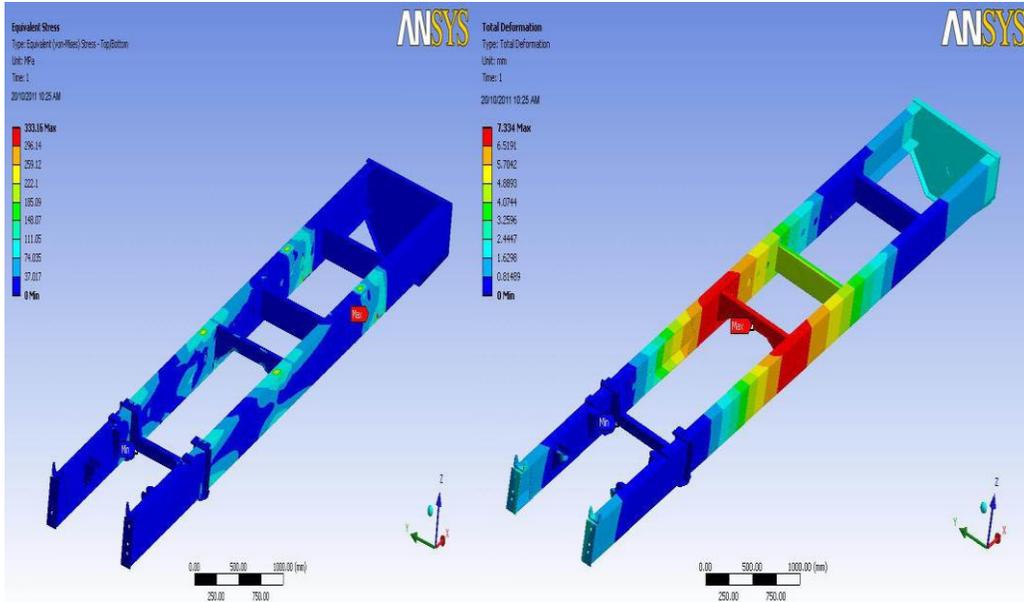
Fig 7 Stress induced in the chassis   Fig 8 Deflection induced in the chassis

Max. Stress = 333.16 Mpa. / Location = just under the loading point
Stress induced in the chassis when the stiffener is 4 mm
Max. Deflection = 7.3 mm / Location = Center portion of rear chassis
Deflection developed in the chassis when the stiffener is 4 mm
        From the above stress and deformation contour, stress induced in the frame is 333.16 Mpa and deformation is 7.3mm. It is is less then the acceptable value. But factor of safety is 1.27. So it is necessary to increase the stiffness thickness of the chassis frame to reduce the stress level.

For stiffener thickness of 5 mm
        Instead of 4mm thickness stiffeners 5mm stiffener was introduced in the chassis frame, static analysis was carried out, and the stress and deflection contours are shown below,
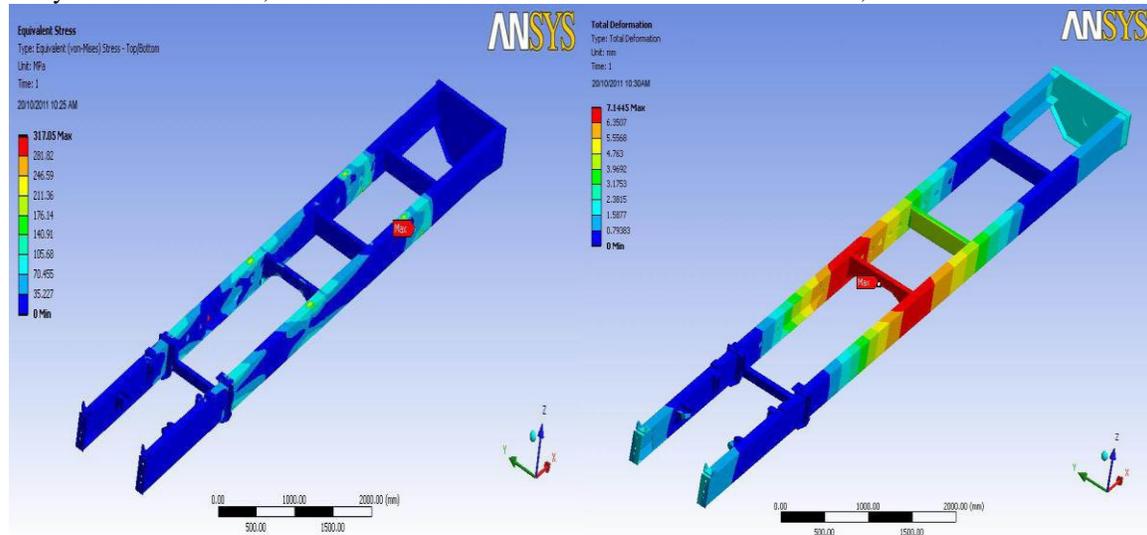
Fig 9 Stress induced in the chassis   Fig10Deflection induced in the chassis

Max. Stress = 317 Mpa. / Location = just under the loading point

Stress induced in the chassis when the stiffener is 5 mm.
Max. Deflection = 7.14 mm / Location = Center portion of rear chassis
Deformation developed in the chassis when the stiffener is 5 mm.

From the above stress and deformation contour, stress induced in the frame is 317 Mpa and deformation is 7.14mm. It is less then the acceptable value. But factor of safety is 1.29. So it is necessary to reinforce the "C" channel with stiffness thickness of the chassis frame to reduce the stress level.

Static Analysis Of Reinforced "C" Channel With Stiffener

Two reinforced "C" channel was introduced under the loading point and the analysis is carried for this condition.
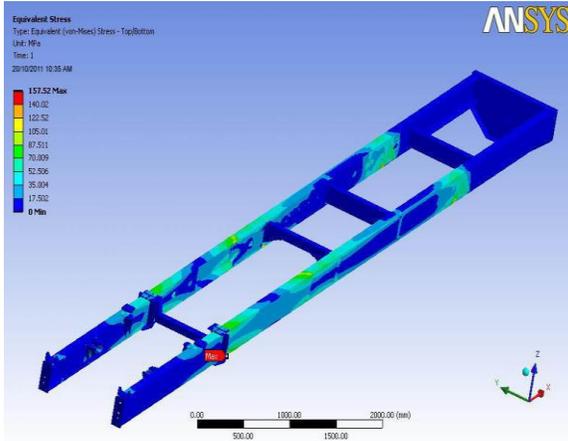
Fig 11Stress induced in the chassis   Fig 12: Deflection induced in the chassis
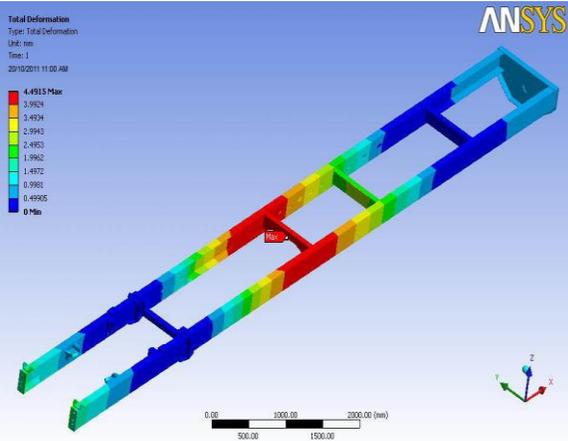
Max. Stress = 157.5 Mpa. / Location = near the supporting point
Max. Deflection = 4.9 mm / Location = Center portion of rear chassis

Factor o Safety
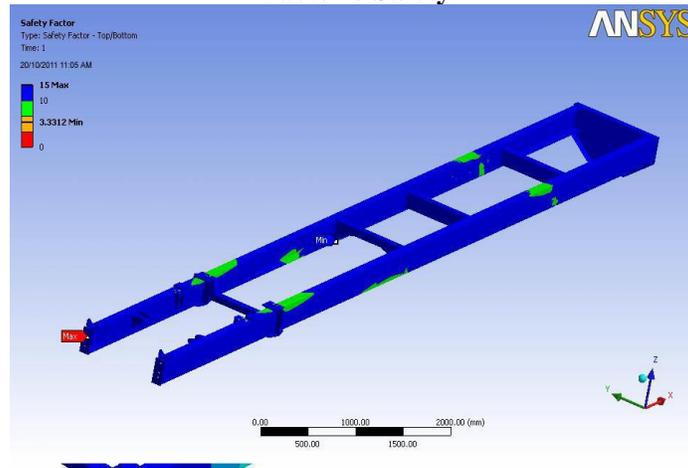
Fig 13: Factor of safety

Dynamic analysis of the chassis

The analysis inertia force is considered to act on z direction and other loading conditions are kept same. Analysis has been carried out in this and plot has shown below.
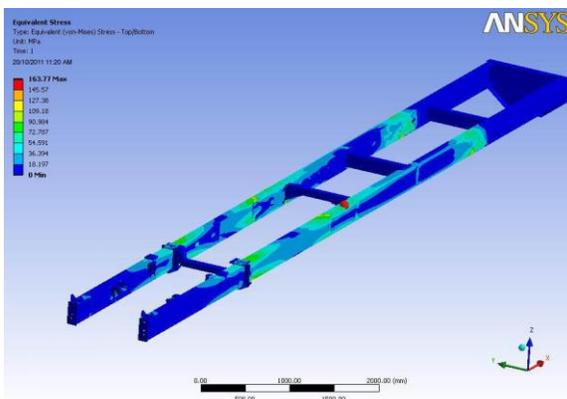
Fig 17: Stress induced in the chassis   Fig 18: Total deformation in the chassis
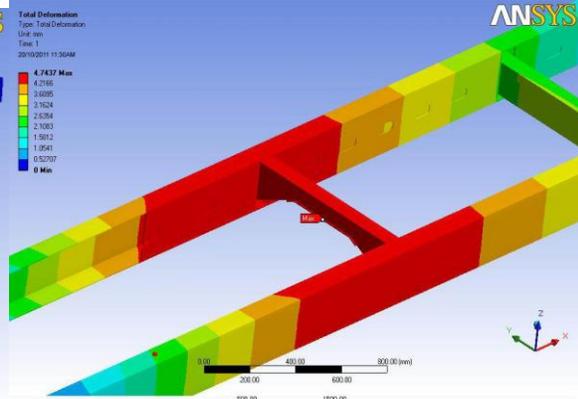
Max. Stress = 163.7 Mpa. / Location = near the supporting point
Max. Deflection = 4.7 mm / Location = Center portion of rear chassis

Factor Of Safety
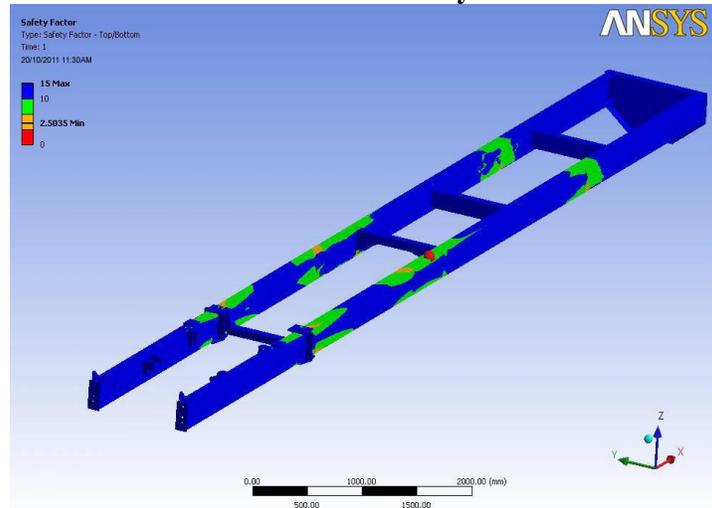
Fig 19:Factor of safety

Static stress Analysis result due to electronic components and Antenna, Engine, Cabin load is shown in Fig .

Dynamic load Stress Analysis result due to Electronic components and Antenna, Engine, Cabin load and inertia load is shown in Fig .

The maximum stress occurs where near the fixed support shown in Fig . The magnitude of maximum stress level is found to be 163.77 N/mm2 which is well within the acceptable criteria.

7.0 CONCLUSIONS

The existing TATRA chassis was analyzed by the finite element analysis for installation of the Antenna and Electronic components and the stress levels are found to be 737.3 N/mm2.After modifications, the TATRA Chassis with suitable reinforcement, increase in thickness, addition of stiffeners, the finite element analysis was carried out, and the stress levels of chassis are found as 173.38 N/mm2, which is less than yield stress 410 N/mm2.

From the above Results, it can be concluded that the modified TATRA chassis is capable to carry the loads beyond the previous payload upto 14 tonnes.

8.0 REFERENCES

1. Sujatha C & V Ramamurti, Bus Vibration Study-Experimental Response to Road Undulation, Int. J. Vehicle Design, Vol. 11, no. 4/5, pp 390-400, 1990.

2. H J Beermann, English translation by Guy Tidbury, The Analysis of Commercial Vehicle Structures, Verlag TUV Rheinland GmbH Koln-1989.

3. Thomas D Gillespie, Fundamentals of Vehicle Dynamics, SAE 1999.

4. J Reimpell & H Stoll, The Automotive Chassis: Engineering Principles, SAE-2000.

5. J Reimpell & H Stoll, The Automotive Chassis: Engineering Principles, ARNOLD-1996.

6. John Fenton, Handbook of Automotive Body Construction & Design Analysis, Professional Engineering Publishing-1998.

7. John Fenton, Handbook of Automotive Body Systems Design, Professional Engineering Publishing-1998.

8. T. R. Chandupatla, A D Belegundu, Introduction to Finite Element in Engineering, PHI-2000.

9. V. Ramamurti, Computer Aided Mechanical Design & Analysis, Tata McGraw Hills-2000.

A Genetic Programming – DWT Hybrid Face Recognition Algorithm

Norolhoda Shahedi ,Mohsen Soryani, Behzad Bozorgtabar

(Department of Electrical Engineering / Iran University of Science and Technology, Iran)

ABSTRACT

Increasing demand for a fast and reliable face recognition technology has obliged researchers to try and examine different pattern recognition schemes. But until now, Genetic Programming (GP), acclaimed pattern recognition, data mining and relation discovery methodology, has been neglected in face recognition literature. This paper tries to apply GP to face recognition. First Discrete Wavelet Transform (DWT) is used to extract features, and then GP is used to classify image groups. To further improve the results, a leveraging method is also utilized. It is shown that although GP might not be efficient in its isolated form, a leveraged GP can offer results comparable to other Face recognition solutions.

I.   INTRODUCTION

Face recognition has become one of the most active research areas of pattern recognition since the early 1990s. In the past 20 years, significant advances have been made in design of successful classifier for face recognition [1]. However the diversity of the face patterns makes it difficult to create robust recognition systems and the complexity of the algorithms makes them hard to implement.

The wavelet transform has many unique features that have made it a popular method for the purpose of image processing and compression. The wavelet transform performs a high degree of decorrelation between neighboring pixels, and it provides a distinct localization of the image in the spatial as well as the frequency domain. This transform also provides an elegant sub-band framework in which both high and low frequency components of the image can be analyzed separately [2]. In this paper we used Discrete Wavelet Transform (DWT) for feature extraction. DWT coefficients are obtained by passing the image through the series of filter bank stages. The procedure of appropriate design of DWT and then selecting the low frequency approximation sub-band leads to improve the robustness of features space with respect to variation in illumination.

Genetic programming is an evolutionary algorithm methodology inspired by biological evolution [3]. Evolutionary algorithms create a population of abstract representations of candidate solutions, which is evolved using biology inspired operators such as selection, cross-over and mutation towards better solutions. In recent years, Genetic Programming and other evolutionary algorithms

has been used in classification and pattern recognition problems [4-5], although to the authors' knowledge, Genetic Programming has never been used in Face Recognition domain.

In many applications, Genetic programming yields simplified symbolical representation of the underlying system it tries to model. This leads to efficient checking of a new sample [6]. On the other hand the complexity and the time needed to find such representation discourages its use in many applications.

Leveraging algorithms are a group of deterministic algorithms where a set of weak leaners are used to create a strong learner [7]. While it is not algorithmically constrained, most leveraging algorithms iteratively employ weak learners based on a distribution and combine them with weighting to form a final strong learner.

In this paper, Genetic Programming is utilized to classify face images which is applied to the extracted features. Using a training group, Genetic Programming discovers possible relationship between the extracted features, which is in turn used to classify new images. To improve results, a leveraging scheme is introduced, which employs Genetic Programming as a weak learner, and combine results of several Genetic Programming classifications as a single strong classifier.

The rest of paper is organized as follows: In section II and III, DWT and Genetic Programming are introduced respectively. Section IV presents the introduced algorithm, where Genetic Programming is used with and without leveraging. In section V, simulations are done on a selected face database and results are compared to previous studies.

II.   DISCRETE WAVELET TRANSFORM

The Discrete Wavelet Transform (DWT) is used for feature extraction. Recall that the wavelet decomposition of an image is done as follows: In the first level of decomposition, the image is split into four sub-bands, namely HH1, HL1, LH1, and LL1, as illustrated in Fig 1. The HH1 sub-band gives the diagonal details of the image; the HL1 sub-band gives the horizontal features, while the LH1 represents the vertical structures. The LL1 sub-band is the low resolution residual consisting of low frequency components and it is this sub-band which is further split at higher levels of decomposition [8]. Fig 2 is an image from ORL Face Database with images obtained one-levelwavelet and after three-level wavelet transform respectively.
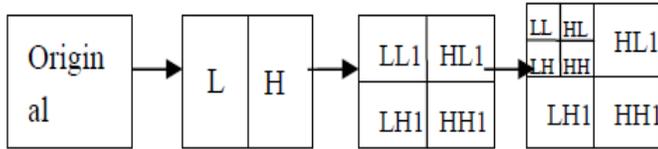
Fig.1. The process of decomposing a face image.
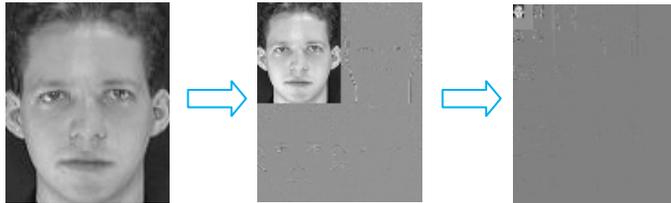
Fig.2. Original face image with figures after one-level and three-level DWT transform respectively.

III. GENETIC PROGRAMMING

Genetic programming is a methodology inspired by biological evolution to find equations, computer programs, analog circuits or in general any suitable structure for a predefined problem [6]. Genetic programming's general mechanisms are almost identical to genetic algorithms, as genetic programming is considered either a specialized form of genetic algorithms or an expansion of it [3]. Genetic programming is usually implemented similar to the following algorithm:

1. Create initial population. Individual solutions (called chromosomes) are usually generated randomly.
2. Evaluate the fitness of each individual in the population.
3. Select best-ranking individuals to reproduce.
4. Breed new generation through crossover and/or mutation (genetic operations) and give birth to offspring.
5. Repeat from step 2 until a termination condition is reached (time limit or sufficient fitness achieved).

Fig. 3 illustrates the general Genetic programming algorithm. Genetic programming's chromosome is traditionally represented by a tree structure, where each node can be function, operator, variable or constant number. Trees can be evaluated in a recursive manner, in which each node's operator or function is executed up on the results of its children's evaluation. Tree structure can easily represent a mathematical equation or a Turing complete program.

IV. CLASSIFICATION ALGORITHM
A. Using Genetic Programming

To classify a given dataset, it is usually enough to find a way for differentiating classes. Using genetic programming, this translates to finding a function which outputs a unique value for each different class:
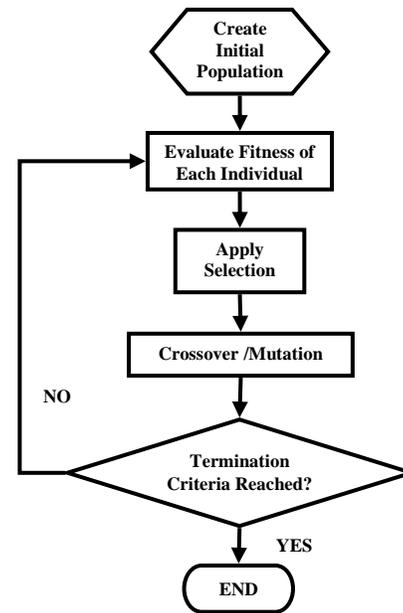
Fig. 3. Genetic programming's flowchart.

$$f(X) = \begin{cases} 0 & X \in C_0 \\ 1 & X \in C_1 \\ \dots \\ n & X \in C_n \end{cases} \qquad (1)$$

This is proven to be difficult. As a result, genetic programming is used to find a function per class that can discriminate only a certain class from others:

$$f_i(X) = \begin{cases} 1 & X \in C_i \\ 0 & X \notin C_i \end{cases} \qquad (2)$$

This method creates N different functions for a total of N classes. Test images are tested one by one against the functions, and the first function to return a non-zero value is used to determine the image's class.

B. Leveraging Algorithm

Leveraging is a method of using multiple results to improve detection. A leveraging algorithm employs multiple weak classifiers to create a strong classifier. The following leveraging algorithm is used in this paper: Instead of using all training images as input, the whole group is partitioned to k different groups. Detector function $f_{i,j}$ is then obtained as a function which can detect class i from other classes in group j. To further improve the results of classification, algorithm above could be repeated N times. For a given image X, the following equation creates the results of classification:

$$C = \sum_{j,n} f_{j,n}(p) \times \frac{1}{1 + \frac{1}{N} err_{j,n}} \qquad (3)$$

where $f_{j,m}$ is result of nth iteration on the jth group, n is the iteration number from total N repetitions, and $err_{j,k}$ is sum of total errors for all images in the training group.

To determine a new image's class, all values acquired from (3) are compared. The class which yields the greatest C is nominated as the new candidate class. It should be noted that a threshold could be defined, as if the results of all classifiers for an image yield lower than a certain value, the image is certainly misclassified.

V. SIMULATION AND RESULTS

The algorithms were implemented in Python and then were tested on the ORL face image database [9]. The ORL database consists of 40 groups, each containing ten 112×92 gray scale images of a single subject. Each subject's images differ in lighting, facial expression and details (smiling/frowning, glasses/no glasses, etc.). Some sample images are displayed in Fig. 4.

Two set of images were created from the ORL database; For the Five-to-Five dataset, five random images of each group were selected for training while the others were used for testing. For the Leave-One-Out set, 9 images were used for training and the remaining image was kept for validation.

First Genetic Programming was tested without leveraging. To evolve the population, an Evolutionary Strategy (ES) of 1+λ with λ = 4 was chosen. Mutation rate was set to 15 percent. The selected function set was {+, -, ×, <, >, MIN, MAX, AND, OR, NOT, CNST} where Boolean operators first compare their operands with 0 and CNST returns a random constant floating point number in range of [-10, 10]. Inputs were chosen from all available DWT features. To limit algorithm time and prevent bloat, each chromosome's depth was limited to 25 and a maximum of 20000 iterations for each evolution was maintained.

To test the leveraged algorithm, algorithm was executed with the same parameters. Also the number of iterations was set to $N = 8$, while the set was divided to $k = 8$ different groups.

Table1 shows a few of discovered relationship functions for a set of pictures. It could be seen that the generated formulas are often simple while only depending on a few components and as a result have a relatively low computational overhead.

Results are brought in Table 2, where they are compared to Euclidean [10] and SVM classifiers. It is observed that Genetic Programming without leveraging has the worst results. On the other hand, Leveraged Genetic programming beats other methods in Five-to-Five. In leave-one-out the results are repeated for Genetic Programming, although this time Leveraged Genetic Programming fell %0.5 (one image in total of 40 images) short of SVM.

Fig. 4. Samples from ORL Face Database

To further investigate Genetic Programming's performance, number of partitioned class groups was changed and the results were brought in Table 3. It was observed that the further partitioning of the images increases recognition error, while decreasing k might mandates increase in time spent for Genetic Programming's evolution.

VI. CONCLUSION

Genetic programming is a general purpose search algorithm that can be utilized in classification problems. In this paper, Genetic programming was exploited to classify face images. The results showed that Genetic Programming alone is not suitable, as required time and computational overhead surpasses that of other methods, and also its recognition ratio is usually lower.

To improve results, a leveraging algorithm was applied to Genetic Programming. The leveraged Genetic Programming in combination with DWT feature extractor showed a good recognition rate, comparable to or in some cases even better than that of other methods.

It is shown that Genetic Programming produces results that usually have a simple structure and therefore a very low computational overhead. Once the system is trained, results can be computed quickly and with lower memory requirements. This might prove lucrative for embedded systems programmers, which have storage and processing constraints.

Table 1

Examples of Acquired Relationship Functions for Detecting Image Group1. *DWT[n]* is The Nth Value on DWT Vector.

No	Function
1	*(DWT[8]- MAX(DWT[15], DWT[7]))> DWT[11]*
2	*AND((DWT[2]< DWT[13]), MAX(DWT[3], 3))*
3	*(DWT[0]× NOT(DWT[5]))*
4	*(DWT[12]× (DWT[20]> (DWT[2]- DWT[18])))*

Table 2

Comparison of Different Algorithms' Recognition Rate

Method	Five-to-Five	Leave One Out
DWT+Euclidean distance	88%	90%
DWT+SVM	91%	93%
DWT+GP	64%	66.5%
DWT+Leveraged GP	92%	92.5%

Table 3

Effect of Number of Partitions in Leveraged Genetic Programming on Recognition Rate

Number of Partitions	Recognition Rate
2	89%
4	92%
5	92%
8	91.5%
10	90%

REFERENCES

[1] S. Liu, Y. Tian, D. Li, New research advances of facial expression recognition, in *International Conf on Machine Learning and Cybernetics,* Baoding, China, *Vol. 2*, July 2009, 1150-1155.

[2] M.Ghazel, Adaptive Fractal and Wavelet Image Denoising, Waterloo, Ontario, Canada, 2004.

[3] J. R. Koza, *Genetic Programming: On the programming of Computer by Means of Natural Selection,* MIT Press: Cambridge, MA, 1992.

[4] S. Xuesong, Y. Zhou, Gray Intensity Images Processing for PD Pattern Recognition Based on Genetic Programming, in *International Joint Conf. on Artificial Intelligence JCAI '09*, Haikou, China, 2009, 711-714.

[5] A. Teredesai, V. Govindaraju, Issues in Evolving GP based Classifiers for a Pattern Recognition Task, in *Proc. 2004 IEEE Congress on Evolutionary Computation*, 20-23 June 2004, 509-515.

[6] J.R. Koza, M.A. Keane, M.J. Streeter, W. Mydlowec, J. Yu and G. Lanza, *Genetic Programming IV: Routine Human-Competitive Machine Intelligence,* Kluwer Academic Publishers, Norwell, MA, 2003.

[7] N. Krause, Y. Singer, Leveraging the margin more carefully, in *Proc. of the twenty-first international conference on Machine learning*, Banff, Alberta, Canada, 2004, 63.

[8] D.L.Donoho, Nonlinear Wavelet Methods for Recovery of Signals, Densities and Spectra from Indirect and Noisy Data, *Proc. Symposia in Applied Mathematics*, oo,1993, 173-205.

[9] ORL, "The database of faces". [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/faced atase.html,2012.

[10] J. Wang, Y. Chen, M. Adjouadi, A comparative study of multilinear principal component analysis for face recognition, *37th IEEE Applied Image Pattern Recognition Workshop*, 2008, 1-6.

A New Article of UPFC Design for Low Frequency Oscillations Using Heuristic Algorithm

A.K. Baliarsingh, D.P.Dash, K.C.Meher

A.K. Baliarsingh Professor in the Electrical Engineering Department, Orissa Engg. College Bhubaneswar, Orissa, India
D.P.Dash is working as Assoc. Professor in the Department of Electrical s Engineering, Orissa Engg. College Bhubaneswar, Orissa, India
K.C.Meher is working as Asst.Professor in Electrical Engineering Department   Orissa Engg. College Bhubaneswar ,Orissa,India

Abstract

This paper presents a systematic approach for designing Unified Power Flow Controller (UPFC) based supplementary damping controllers for damping low frequency oscillations in a single-machine infinite-bus power system. Detailed investigations have been carried out considering the four alternatives UPFC based damping controller namely modulating index of series inverter (m_B), modulating index of shunt inverter (m_E), phase angle of series inverter (δ_B) and phase angle of the shunt inverter (δ_E). The proposed controllers is formulated as an optimization problem and Heuristic Optimization Method (HOM) is employed to optimize damping controller parameters. Simulation results are presented and compared with a conventional method of tuning the damping controller parameters to show the effectiveness and robustness of the proposed design approach.

Keywords—Power System Oscillations, Heuristic Optimization (HO), Flexible AC Transmission Systems (FACTS), Unified Power Flow Controller (UPFC), Damping Controller.

I. INTRODUCTION

THE main causes of the power systems to be operated near their stability limits is due to the fact that power systems are today much more loaded than before as power demand grows rapidly and expansion in transmission and generation is restricted with the limited availability of resources and the strict environmental constraints. In few occasions interconnection between remotely located power systems gives rise to low frequency oscillations in the range of 0.2-3.0 Hz. These low frequency oscillations are also observed when large power systems are interconnected by relatively weak tie lines. If the system is not well damped, these oscillations may keep increasing in magnitude until loss of synchronism results [1]. The installation of Power System Stabilizer (PSS) is both economical and effective; in order to damp these power system oscillations and increase system oscillations stability. During the last decade, continuous and fast improvement of power electronics technology has made Flexible AC Transmission Systems (FACTS) a promising concept for power system applications [2-4]. With the application of FACTS technology, power flow along transmission lines can be

more flexibly controlled. Due to the fact of the extremely fast control action is associated with FACTS-device operations, they have been very promising candidates for utilization in power system damping enhancement. The Unified Power Flow Controller (UPFC) is regarded as one of the most versatile devices in the FACTS device family [5-6] which has the capability to control of the power flow in the transmission line, improve the transient stability, alleviate system oscillation and offer voltage support. UPFC can provide simultaneous and independent control of important power system parameters: line active power flow, line reactive power flow, impedance; and voltage. In that way, it offers the essential functional flexibility for the collective application of phase angle control with controlled series and shunt compensation [2].

A conventional lead-lag controller structure is preferred by the power system utilities because of the ease of on-line tuning and also lack of assurance of the stability by some adaptive or variable structure techniques [7-10]. Traditionally, for the small signal stability studies of a power system**,** the linear model of Phillips-Heffron has been used for years, providing reliable results [1]. Although the model is a linear model, it is quite accurate for studying low frequency oscillations and stability of power systems [11-12]. The problem of UPFC damping controller parameter tuning is a complex exercise. A number of conventional techniques have been reported in the literature pertaining to design problems of conventional power system stabilizers namely the pole placement technique [13], phase compensation/root locus technique [14-15], residue compensation [16], and also the modern control theory. Unfortunately, the conventional techniques are time consuming as they are iterative and require heavy computation burden and slow convergence. In addition, the search process is susceptible to be trapped in local minima and the solution obtained may not be optimal. Also, the designed controller should provide some degree of robustness to the variations loading conditions, and configurations as the machine parameters change with operating conditions. A set of controller parameters which stabilise the system under a certain operating condition may no longer yield satisfactory results when there is a drastic change in power system operating conditions and configurations [12].

In recent years, one of the most promising research fields has been "Evolutionary Techniques", an area utilizing

analogies with nature or social systems. These techniques constitute an approach to search for the optimum solutions via some form of directed random search process. Evolutionary techniques are finding popularity within research community as design tools and problem solvers because of their versatility and ability to optimize in complex multimodal search spaces applied to non-differentiable objective functions.

Recently, Real Coded GA (RCGA) is appeared as a promising Swarm technique for handling the optimization problems [17]. It has been popular in academia and the industry mainly because of its intuitiveness, ease of implementation, and the ability to effectively solve highly nonlinear, mixed integer optimisation problems that are typical of complex engineering systems. It has been reported in the literature that RCGA is more efficient in terms of CPU time and offers higher precision with more consistent results [8, 18-21]. In view of the above, this paper proposes to use RCGA optimization technique for the damping controller design. For the proposed controller design, a time-domain based employing integral of time multiplied by speed deviation error has been employed. The optimal UPFC controller parameters are obtained employing PSO. The proposed damping controllers are tested on a weakly connected power system with different disturbances with parameter variations. Simulation results

are presented and compared with a conventional tuning technique to show the effectiveness and robustness of the proposed approach.

The reminder of the paper is organized in five major sections. Power system modeling with the proposed UPFC-based supplementary damping controller is presented in Section II. The design problem and the objective function are presented in section III. In Section IV, an overview of PSO is presented. The results are presented and discussed in Section V. Finally, in Section VI conclusions are given.

II. MODELING THE POWER SYSTEM WITH UPFC DAMPING CONTROLLER

The single-machine infinite-bus (SMIB) power system installed with a UPFC as shown in Fig. 1 is considered in this study. The UPFC is installed in one of the two parallel transmission lines. This arrangement, comprising two parallel transmission lines, permits the control of real and reactive power flow through a line. The static excitation system, model type IEEE-STIA, has been considered. The UPFC is assumed to be based, on pulse width modulation (PWM) converters. The nominal loading condition and system parameters are even in Appendix 1.
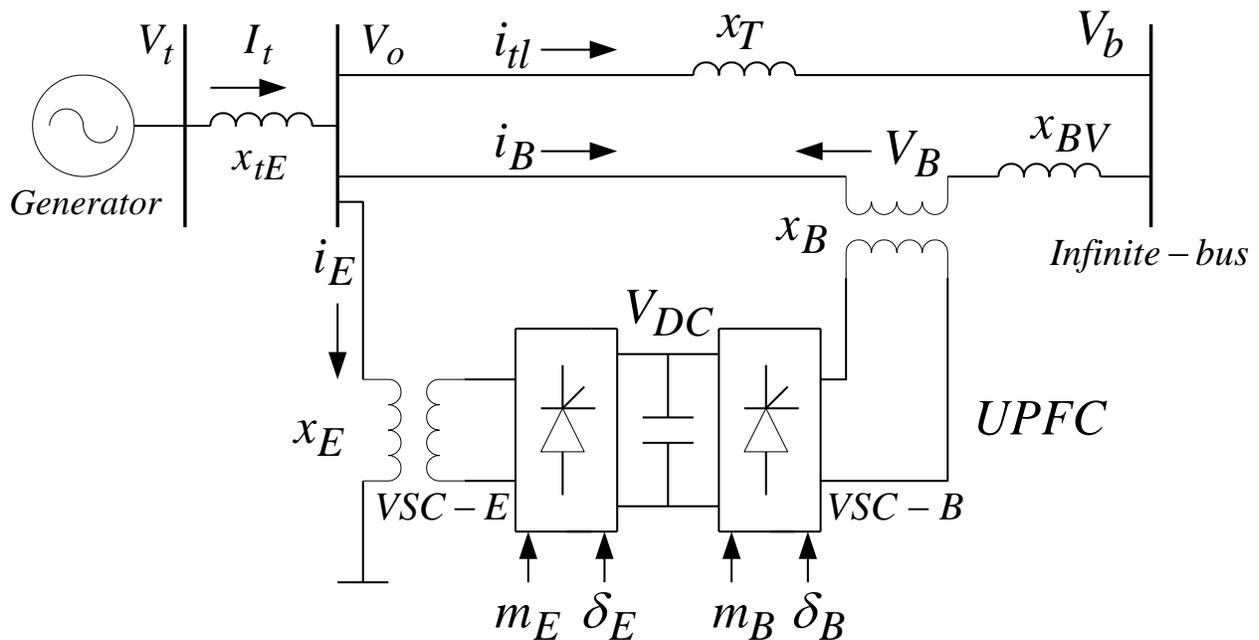
Fig. 1. Single-machine infinite-bus power system with UPFC

A. Non-Linear Equations

The non-linear differential equations of the SMIB system with UPFC is obtained by neglecting the resistances of the components of the system (i.e. generator, transformer, transmission lines, and shunt and series converter transformers) and the transient associated with the stator of the synchronous generator, transmission lines and transformers of

the UPFC. The nonlinear dynamic model of the system with UPFC is given below [22-23]:

$$\dot{\delta} = \omega_{\circ}(\omega - 1) \tag{1}$$

$$\dot{\omega} = \frac{(P_m - P_e - D\Delta\omega)}{M} \qquad (2)$$

$$\dot{E}_q' = \frac{1}{T_{do}'}[-E_q + E_{fd}] \qquad (3)$$

$$\dot{E}_{fd}' = \frac{K_A}{1 + sT_A}[V_{ref} - V_t] \qquad (4)$$

$$V_{dc} = \frac{3m_E}{4C_{dc}}\left(I_{Ed}Sin\delta_E + I_{Eq}Cos\delta_E\right)$$
$$+ \frac{3m_B}{4C_{dc}}\left(I_{Bd}Sin\delta_B + I_{Bq}Cos\delta_B\right) \qquad (5)$$

where,

$$P_e = V_{td}I_{td} + V_{tq}I_{tq} \qquad (6)$$

$$E_q = E_q' + \left(X_d - X_d'\right)I_{td} \qquad (7)$$

$$V_t = V_{td} + jV_{tq} \qquad (8)$$

$$V_t = X_q I_{tq} \qquad (9)$$

$$V_{tq} = E_q' - X_d' I_{td} \qquad (10)$$

$$I_{td} = I_{tld} + I_{Ed} + I_{Bd} \qquad (11)$$

$$I_{tld} = \frac{X_E}{X_T}I_{Ed} + \frac{1}{X_T}\frac{m_E V_{dc}}{2}Cos\delta_E$$
$$- \frac{1}{X_T}V_b Cos\delta \qquad (12)$$

$$I_{tlq} = \frac{X_E}{X_T}I_{Eq} - \frac{1}{X_T}\frac{m_E V_{dc}}{2}Sin\delta_E$$
$$+ \frac{1}{X_T}V_b Sin\delta \qquad (13)$$

$$I_{Ed} = \frac{(X_{dT} + X_{BB}X_{b3})}{X_{dE}}V_b Cos\delta$$
$$- \frac{(X_{dT} + X_{BB}X_{b2})}{X_{dE}}\frac{m_E V_{dc}}{2}Cos\delta_E \qquad (14)$$
$$+ \frac{X_{BB}}{X_{dE}}E_q' - \frac{X_{dT}}{X_{dE}}\frac{m_B V_{dc}}{2}Cos\delta_B$$

$$I_{Eq} = \frac{(X_{qT} + X_{BB}X_{a3})}{X_{qE}}V_b Sin\delta$$
$$- \frac{(X_{qT} + X_{BB}X_{a2})}{X_{qE}}\frac{m_E V_{dc}}{2}Sin\delta_E \qquad (15)$$
$$- \frac{X_{qT}}{X_{qE}}\frac{m_B V_{dc}}{2}Sin\delta_B$$

$$I_{Bd} = \frac{1}{X_{dE}}\left(X_E E_q' + (X_{b1} - X_E X_{b2})\right)$$
$$\frac{m_E V_{dc}}{2}Cos\delta_E + \frac{(X_{b3}X_E - X_{b1})}{X_{qE}}V_b Cos\delta \qquad (16)$$
$$+ X_{b1}\frac{m_B V_{dc}}{2}Cos\delta_B$$

$$I_{Bq} = \frac{1}{X_{qE}}\begin{pmatrix} (X_{a1} - X_E X_{a2})\dfrac{m_E V_{dc}}{2}Sin\delta_E \\ + \dfrac{(X_{a3}X_E - X_{a1})}{X_{qE}}V_b Sin\delta \\ + X_{a1}\dfrac{m_B V_{dc}}{2}Sin\delta_B \end{pmatrix} \qquad (17)$$

The variables used in the above equations are defined as:

$$X_{qT} = X_q + X_{tE} \; ; \; X_{ds} = X_{tE} + X_d' + X_E \; ;$$

$$X_{qs} = X_q + X_{tE} + X_E \; ;$$

$$X_{a1} = \frac{(X_{qs}X_T + X_{qT}X_E)}{X_T} \; ; \; X_{a2} = 1 + \frac{X_{qT}}{X_T} \; ;$$

$$X_{a3} = -\frac{X_{qT}}{X_T} \; ; \; X_{b1} = \frac{(X_{ds}X_T + X_{dT}X_E)}{X_T} \; ;$$

$$X_{b2} = 1 + \frac{X_{dT}}{X_T} \; ; \; X_{b3} = \frac{X_{dT}}{X_T}$$

The equation for the real power balance between the series and shunt converters is

$$R_e\left(\bar{V}_B\bar{I}_B^* + \bar{V}_E\bar{I}_E^*\right) = 0 \qquad (18)$$

B. Linearized Equations

In the design of electromechanical mode damping stabilizer, a linearized incremental model around an operating point is usually employed. The Phillips-Heffron model of the power system with FACTS devices is obtained by linearizing the set of equations (1) around an operating condition of the power system. The linearized expressions are as follows [22-23]:

$$\Delta \dot{\delta} = \omega_0 \Delta \omega \tag{19}$$

$$\Delta \dot{\omega} = \frac{(\Delta P_m - \Delta P_e - D\Delta\omega)}{M} \tag{20}$$

$$\Delta \dot{E}' = \frac{\left(-\Delta E_q + \Delta E_{fd}\right)}{T_{do}'} \tag{21}$$

$$\Delta \dot{E}_{fd} = \frac{-\Delta E_{fd} + K_A\left(\Delta V_{ref} - \Delta V_t\right)}{T_A} \tag{22}$$

$$\Delta \dot{V}_{dc} = K_7\Delta\delta + K_8\Delta E_q' - K_9\Delta V_{dc}$$
$$+ K_{ce}\Delta m_E + K_{c\delta e}\Delta\delta_E + K_{cb}\Delta m_B + K_{c\delta b}\Delta\delta_B \tag{23}$$

where,

$$\Delta P_e = K_1\Delta\delta + K_2\Delta E_q' + K_{pe}\Delta m_E$$
$$+ K_{p\delta e}\Delta\delta_E + K_{pb}\Delta m_B + K_{p\delta b}\Delta\delta_B + K_{pd}\Delta V_{dc}$$

$$\Delta E_q = K_4\Delta\delta + K_3\Delta E_q' + K_{qe}\Delta m_E$$
$$+ K_{q\delta e}\Delta\delta_E + K_{qb}\Delta m_B + K_{q\delta b}\Delta\delta_B + K_{qd}\Delta V_{dc}$$

$$\Delta V_t = K_5\Delta\delta + K_6\Delta E_q' + K_{ve}\Delta m_E$$
$$+ K_{v\delta e}\Delta\delta_E + K_{vb}\Delta m_B + K_{v\delta b}\Delta\delta_B + K_{vd}\Delta V_{dc}$$

The modified Phillips-Heffron model of the single-machine infinite-bus (SMIB) power system with UPFC-based damping controller is obtained using linearized equation set (2). The corresponding block diagram model is shown in Fig. 2. The modified Heffron-Phillips model has 28constants as compared to 6 constants in the original Heffron-Phillips model of the SMIB system. These constants are functions of the system parameters and initial operating condition.

By controlling m_B, the magnitude of series injected voltage can be controlled, by controlling δ_B, the phase angle of series injected voltage can be controlled, by controlling m_E, the output voltage of the shunt converter can be controlled and by controlling δ_E, the phase angle of output voltage of the shunt converter can be controlled. The series and shunt converter are

controlled in a coordinated manner to ensure real power output of the shunt converter is equal to the real power input to the series converter. The constancy of the DC voltage ensures that this equality is maintained.

In Fig 2, the row vectors $[K_{pu}], [K_{qu}], [K_{vu}]$ and $[K_{cu}]$ are defined as:

$$[K_{pu}] = [K_{pc} \quad K_{p\delta e} \quad K_{pb} \quad K_{p\delta b}] \tag{24}$$

$$[K_{qu}] = [K_{qe} \quad K_{q\delta e} \quad K_{qb} \quad K_{q\delta b}] \tag{25}$$

$$[K_{vu}] = [K_{ve} \quad K_{v\delta e} \quad K_{vb} \quad K_{v\delta b}] \tag{26}$$

$$[K_{cu}] = [K_{ce} \quad K_{c\delta e} \quad K_{cb} \quad K_{c\delta b}] \tag{27}$$

The control vector $[\Delta u]$ is the column vector defined as follows:

$$[\Delta u] = \begin{bmatrix} \Delta m_E & \Delta\delta_E & \Delta m_B & \Delta\delta_B \end{bmatrix}^T$$

where,

Δm_B - Deviation in modulation index m_B of series converter.

$\Delta\delta_B$ - Deviation in phase angle of the injected voltage.

Δm_E - Deviation in modulation index m_E of shunt converter.

$\Delta\delta_E$ - Deviation in phase angle of the shunt converter voltage.

III. THE PROPOSED APPROACH

A. Structure of UPFC-based Damping Controller

The commonly used lead–lag structure is chosen in this study as UPFC-based supplementary damping controller as shown in Fig. 3. The structure consists of a gain block; a signal washout block and two-stage phase compensation block. The phase compensation block provides the appropriate phase-lead characteristics to compensate for the phase lag between input and the output signals. The signal washout block serves as a high-pass filter which allows signals associated with oscillations in input signal to pass unchanged. Without it steady changes in input would modify the output. The input signal of the proposed UPFC-based controller is the speed deviation $\Delta\omega$ and the output is the change in control vector $[\Delta u]$. From the viewpoint of the washout function the value of washout time constant is not critical in lead-lag structured controllers and may be in the range 1 to 20 seconds [1].
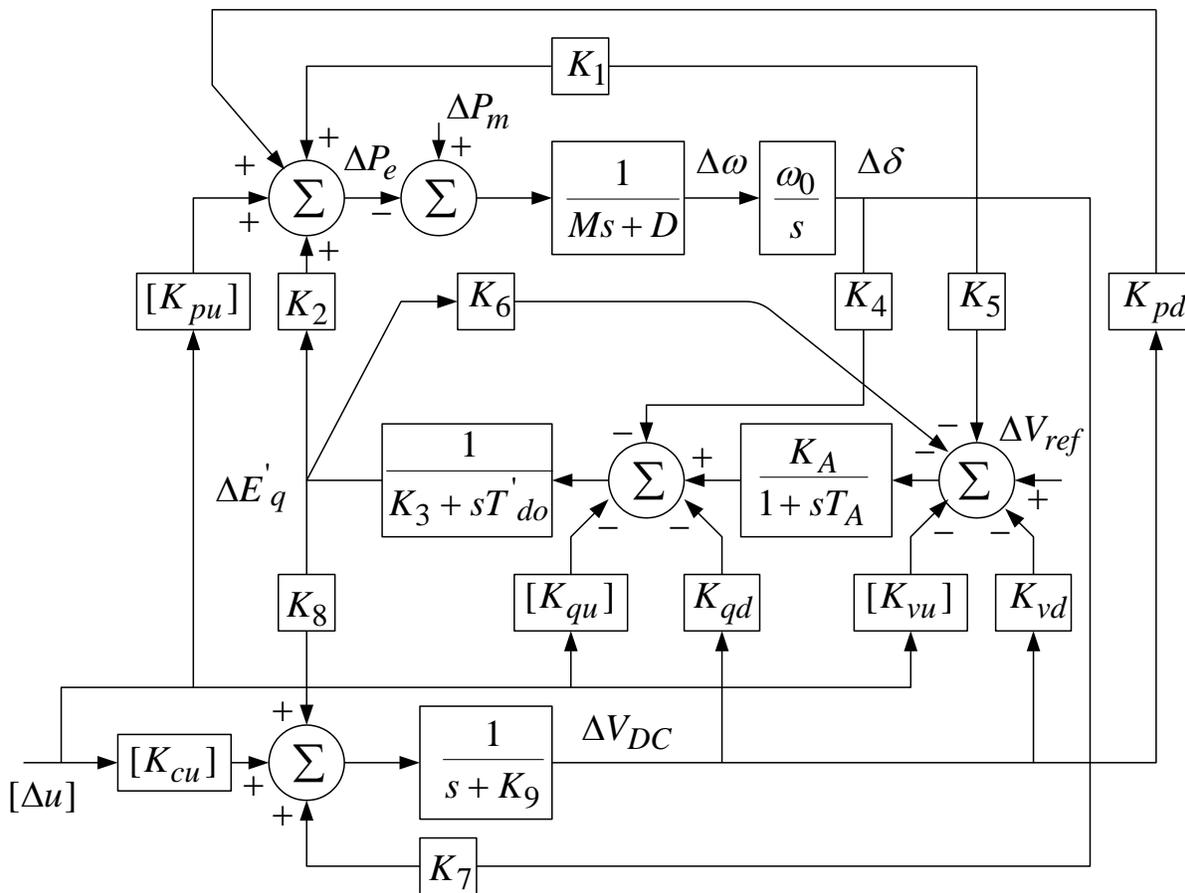
Fig. 2 Modified Heffron-Phillips model of SMIB system with UPFC
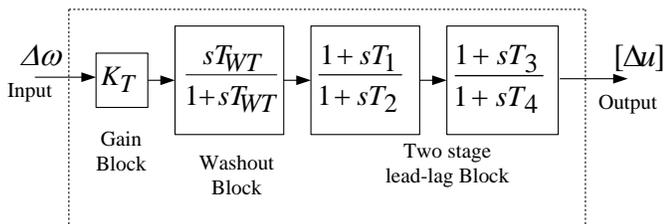
Fig. 3. Structure of the UPFC-based damping controller

From the viewpoint of the washout function the value of washout time constant is not critical in lead-lag structured controllers and may be in the range 1 to 20 seconds [1]. In the present study, a washout time constant of $K_{WT} = 10$ s is used.

The controller gains K_T; and the time constants T_1, T_2, T_3 and T_4 are to be determined.

B. Objective Function

Tuning a controller parameter can be viewed as an optimization problem in multi-modal space as many settings of the controller could be yielding good performance. Traditional method of tuning doesn't guarantee optimal parameters and in most cases the tuned parameters need improvement through trial and error. The aim of any evolutionary optimization technique is basically to optimize

(minimize/maximize) an objective function or fitness function satisfying the constraints of either state or control variable or both depending upon the requirement. It is worth mentioning that the UPFC-based controllers are designed to minimize the power system oscillations after a disturbance so as to improve the stability. These oscillations are reflected in the deviation in the generator rotor speed ($\Delta\omega$). In the present study, an integral time absolute error of the speed deviations is taken as the objective function J, expressed as:

$$J = \int_0^{t_1} t\,|\,e(t)|\,dt \qquad (28)$$

Where, 'e' is the error signal ($\Delta\omega$) and t_1 is the time range of simulation. The parameters of the damping controller are obtained using PSO. A brief overview of PSO is presented in the next section.

IV. OVERVIEW OF HEURISTIC OPTIMIZATION METHOD

Real Coded Genetic Algorithm (RCGA) can be viewed as a general-purpose search method, an optimization method, or a learning mechanism, based loosely on Darwinian principles of biological evolution, reproduction and "the survival of the fittest." GA maintains a set of candidate solutions called population and repeatedly modifies them. At each step, the

GA selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Candidate solutions are usually represented as strings of fixed length, called chromosomes.

Given a random initial population GA operates in cycles called generations, as follows [13]:

- Each member of the population is evaluated using an objective function or fitness function.
- The population undergoes reproduction in a number of iterations. One or more parents are chosen stochastically, but strings with higher fitness values have higher probability of contributing an offspring.
- Genetic operators, such as crossover and mutation, are applied to parents to produce offspring.
- The offspring are inserted into the population and the process is repeated.

Over successive generations, the population "evolves" toward an optimal solution. GA can be applied to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, nondifferentiable, stochastic, or highly nonlinear. GA has been used to solve difficult engineering problems that are complex and difficult to solve by conventional optimization methods.

Implementation of GA requires the determination of six fundamental issues: chromosome representation, selection function, the genetic operators, initialization, termination and evaluation function. Brief descriptions about these issues are provided in the following sections [8, 18-21].

A. Chromosome representation

Chromosome representation scheme determines how the problem is structured in the GA and also determines the genetic operators that are used. Each individual or chromosome is made up of a sequence of genes. Various types of representations of an individual or chromosome are: binary digits, floating point numbers, integers, real values, matrices, etc. Generally natural representations are more efficient and produce better solutions. Real-coded representation is more efficient in terms of CPU time and offers higher precision with more consistent results.

B. Selection function

To produce successive generations, selection of individuals plays a very significant role in a genetic algorithm. The selection function determines which of the individuals will survive and move on to the next generation. A probabilistic selection is performed based upon the individual's fitness such that the superior individuals have more chance of being selected. There are several schemes for the selection process: roulette wheel selection and its extensions, scaling techniques, tournament, normal geometric, elitist models and ranking methods.

The selection approach assigns a probability of selection P_j to each individuals based on its fitness value. In the present study, normalized geometric selection function has been used.

In normalized geometric ranking, the probability of selecting an individual P_i is defined as:

$$Pi = q'(1-q)^{r-1} \qquad (29)$$

$$q' = \frac{q}{1-(1-q)^P} \qquad (30)$$

where,

q = probability of selecting the best individual

r = rank of the individual (with best equals 1)

P = population size

C. Genetic operators

The basic search mechanism of the GA is provided by the genetic operators. There are two basic types of operators: crossover and mutation. These operators are used to produce new solutions based on existing solutions in the population. Crossover takes two individuals to be parents and produces two new individuals while mutation alters one individual to produce a single new solution. The following genetic operators are usually employed: simple crossover, arithmetic crossover and heuristic crossover as crossover operator and uniform mutation, non-uniform mutation, multi-non-uniform mutation, boundary mutation as mutation operator. Arithmetic crossover and non-uniform mutation are employed in the present study as genetic operators. Crossover generates a random number r from a uniform distribution from 1 to m and creates two new individuals by using equations:

$$x_i' = \begin{cases} x_i, & \text{if } i < r \\ y_i & \text{otherwise} \end{cases} \qquad (31)$$

$$y_i' = \begin{cases} y_i, & \text{if } i < r \\ x_i & \text{otherwise} \end{cases} \qquad (32)$$

Arithmetic crossover produces two complimentary linear combinations of the parents, where r = U (0, 1):

$$\bar{X}' = r\,\bar{X} + (1-r)\,\bar{Y} \qquad (33)$$

$$\bar{Y}' = r\,\bar{Y} + (1-r)\,\bar{X} \qquad (34)$$

Non-uniform mutation randomly selects one variable j and sets it equal to a non-uniform random number.

$$x_i' = \begin{cases} x_i + (b_i - x_i)\,f(G) & \text{if } r_1 < 0.5, \\ x_i + (x_i + a_i)\,f(G) & \text{if } r_1 \geq 0.5, \\ x_i, & \text{otherwise} \end{cases} \qquad (35)$$

where,

$$f(G) = \left(r_2\left(1 - \frac{G}{G_{max}}\right)\right)^b \qquad (36)$$

r_1, r_2 = uniform random nos. between 0 to 1.

G = current generation.

G_{max} = maximum no. of generations.

b = shape parameter.

D. Initialization, termination and evaluation function

An initial population is needed to start the genetic algorithm procedure. The initial population can be randomly generated or can be taken from other methods.

GA moves from generation to generation until a stopping criterion is met. The stopping criterion could be maximum number of generations, population convergence criteria, lack of improvement in the best solution over a specified number of generations or target value for the objective function.

Evaluation functions or objective functions of many forms can be used in a GA so that the function can map the population into a partially ordered set.

V. RESULTS AND DISCUSSIONS

A. Application of RCGA

The optimization of the proposed UPFC-based supplementary damping controller parameters is carried out by minimizing the fitness given in equation (28) employing RCGA. For the implementation of RCGA normal geometric selection is employed which is a ranking selection function based on the normalized geometric distribution. Arithmetic crossover takes two parents and performs an interpolation along the line formed by the two parents. Non uniform mutation changes one of the parameters of the parent based on a non-uniform probability distribution. This Gaussian distribution starts wide, and narrows to a point distribution as the current generation approaches the maximum generation.

The model of the system under study has been developed in MATLAB/SIMULINK environment and RCGA programme has been written (in .mfile). For objective function calculation, the developed model is simulated in a separate programme (by .m file using initial population/controller parameters) considering a severe disturbance. Form the SIMULINK model the objective function value is evaluated and moved to workspace. The process is repeated for each individual in the population. For objective function calculation, a 10% increase in mechanical power input is considered. Using the objective function values, the population is modified by RCGA for the next generation. The flow chart of proposed optimization algorithm is shown in Fig. 4.

For different problems, it is possible that the same parameters for GA do not give the best solution and so these can be changed according to the situation. One more important point that affects the optimal solution more or less is the range for unknowns. For the very first execution of the program, more wide solution space can be given and after getting the solution one can shorten the solution space nearer to the values obtained in the previous iteration. The parameters employed for the implementations of RCGA in the present

study are given in Table I. Optimization were performed with the total number of generations set to 100. The optimization processes is run 20 times for both the control signals and best among the 20 runs are provided in the Table II.
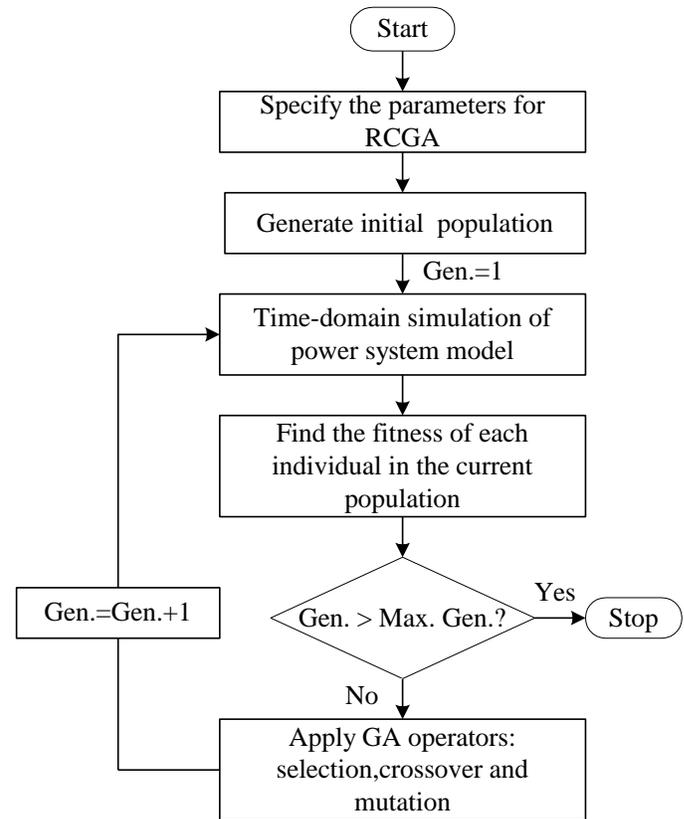
Fig. 4 Flowchart of RCGA optimization process to optimally tune the controller parameters

Table I: Parameters used in RCGA

Parameter	Value/Type
Maximum generations	100
Population size	50
Type of selection	Normal geometric [0 0.08]
Type of crossover	Arithmetic [2]
Type of mutation	Nonuniform [2 100 3]
Termination method	Maximum generation

Table II: Optimized UPFC-based damping controller parameters.

Damping controller	K_T	T_1	T_2	T_3	T_4
m_B - based	89.3312	0.2774	0.3217	0.3294	0.3538
δ_B - based	34.1934	0.1650	0.1173	0.1385	0.3603
m_E - based	19.2086	0.5494	0.4874	0.3656	0.3269
δ_E - based	29.0276	0.1090	0.2463	0.2416	0.2367

B. Simulation Results

To assess the effectiveness and robustness of the proposed damping controllers various disturbances and parameter variations are considered. The performance of the proposed controllers is compared with a published [15] conventional design technique (phase compensation technique). The response with out controller is shown in dashed lines (with legend "WC") and the response with conventional phase compensation technique tuned UPFC-based damping controller is shown dotted lines (with legend "PCT"). The responses with proposed RCGA optimized UPFC-based damping controller are shown in solid lines (with legend 'RCGA').

Case I: m_B - based UPFC damping controller

A 10% step increase in mechanical power input at t = 1.0 s is assumed. The system speed and electrical power deviation response for the above contingency are shown in Figs. 5-6. It is clear from Figs. that without control the system is oscillatory and becomes unstable. Stability of the system is maintained and power system oscillations are effectively damped out with m_B - based UPFC damping controller. It can also be seen from Figs. that the performance of the system is better with the proposed RCGA optimized damping controller compared to the conventionally designed controller.
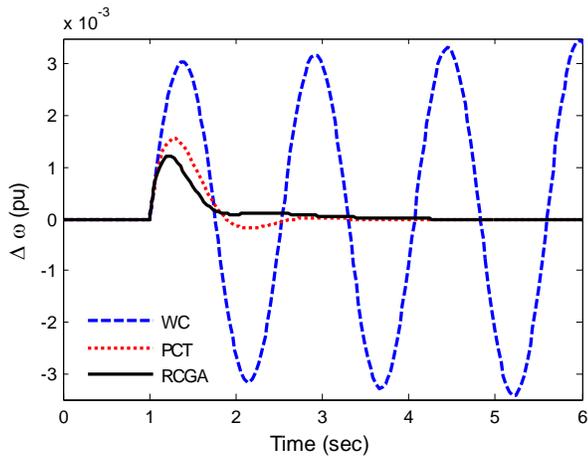
Fig. 5. Speed deviation response for Case-I
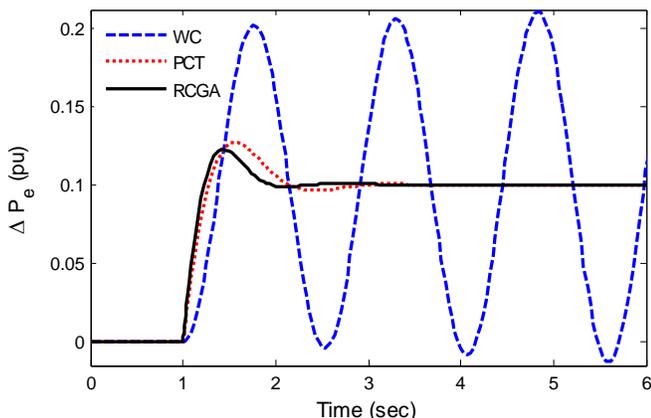
Fig. 6. Electrical power deviation response for Case-I

Case II: δ_B - based UPFC damping controller

The performance of the system for the under same contingency (10% step increase in mechanical power input at t = 1.0 s) is verified with δ_B - based UPFC damping controller and the system response is shown in Figs. 7-8. It can be observed from Figs. that the performance of the system is better with the proposed RCGA optimized damping controller compared to the conventionally designed controller.
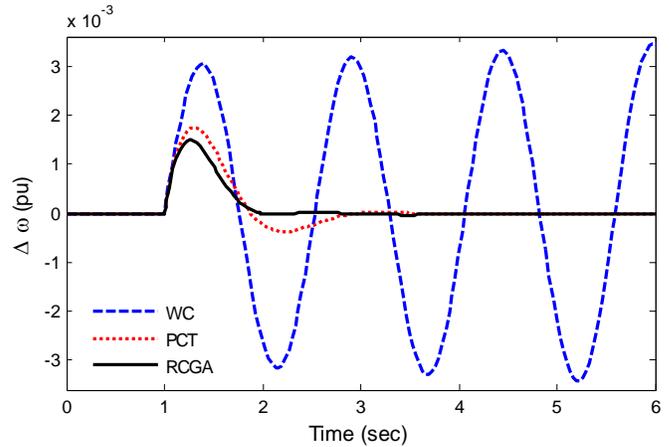
Fig. 7. Speed deviation response for Case-II
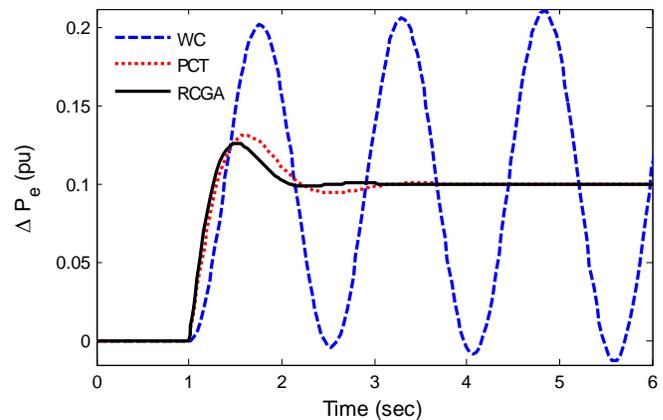
Fig. 8. Electrical power deviation response for Case-II

Case III: m_E - based UPFC damping controller

The performance of the system for the under same contingency (10% step increase in mechanical power input at t = 1.0 s) is demonstrated with m_E - based UPFC damping controller. It can be observed from the system response shown in Figs. 9-10 that the performance of the system is slightly better with the proposed RCGA optimized m_E - based UPFC damping controller compared to the conventional phase compensation technique based designed of m_E - based UPFC damping controller.
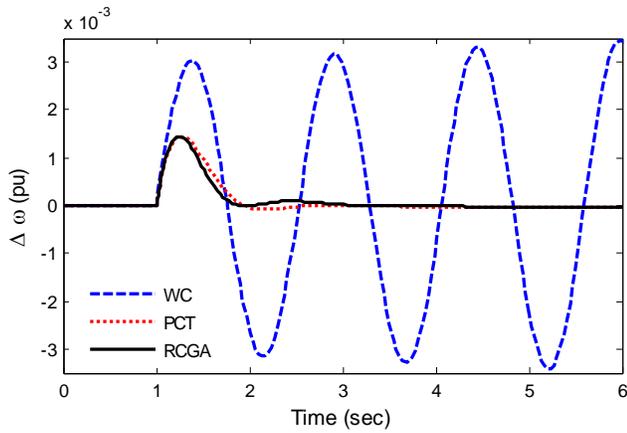
Fig. 9. Speed deviation response for Case-III
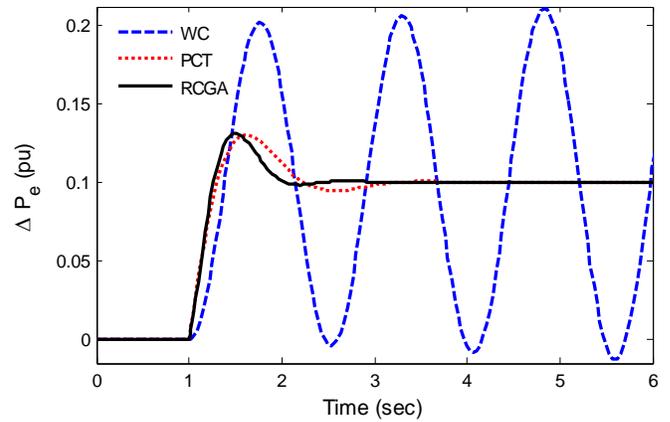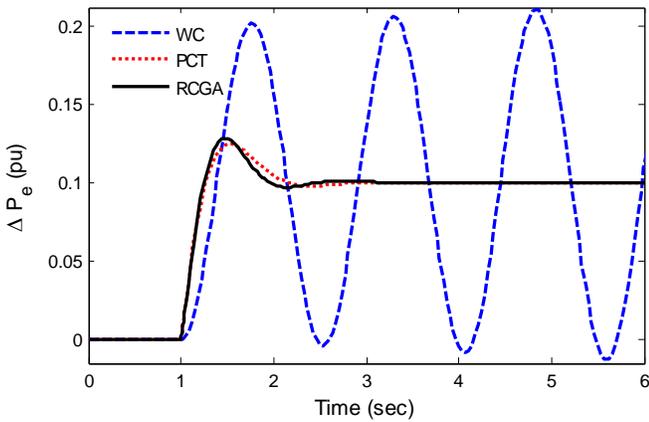
Fig. 10. Electrical power deviation response for Case-III

Case IV: δ_E - based UPFC damping controller

Figs. 11-12 show the system response for the same contingency with δ_E - based UPFC damping controller from which it can be seen that the proposed RCGA optimized δ_E - based UPFC damping controller performs better than the phase compensation technique tuned δ_E - based UPFC damping controller.

Fig. 11. Speed deviation response for Case-IV
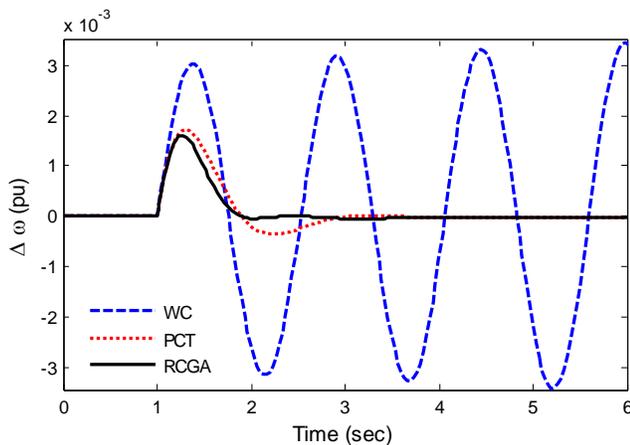
Fig. 12. Electrical power deviation response for Case-IV

Case V: Comparison four alternative UPFC-based damping controllers

Figs. 13-14 shows the system dynamic response considering a step load increase of 10 % and step load decrease of 5% respectively. It can be concluded from the Figs. that all four alternative damping controllers provide satisfactory damping performance for both increase and decrease in mechanical power input. However, the performance of m_B - based UPFC damping controller seems to be slightly better among the four alternatives.
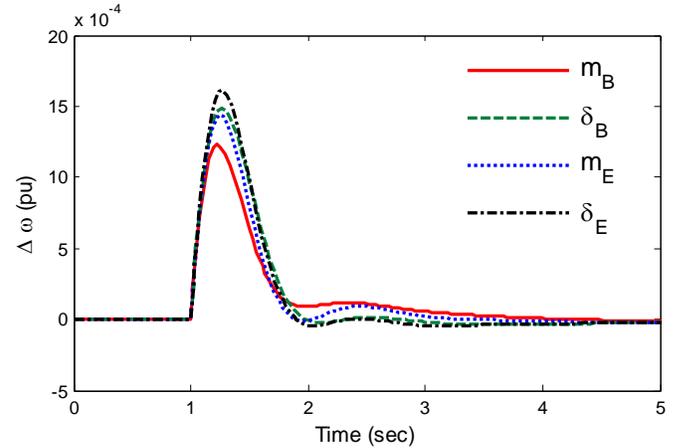
Fig. 13. Speed deviation response for Case-V (step increase in Pm)
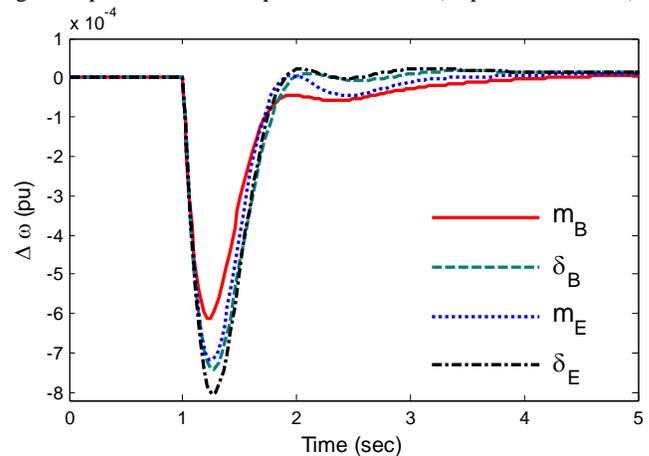
Fig. 14. Speed deviation response for Case-V (step decrease in Pm)

Case VI: Comparison four alternative UPFC-based damping controllers for step change in reference voltage

To test the robustness of the proposed approach, another disturbance is considered. The reference voltage is increased by 5 % at t=1.0 and the system dynamic response with all four alternative damping controllers is shown in Fig. 15. It can be concluded from Fig. 15 that though all four alternative damping controllers provide satisfactory damping performance for the above contingency, the performance of m_B - based UPFC damping controller is slightly better among the four alternatives.
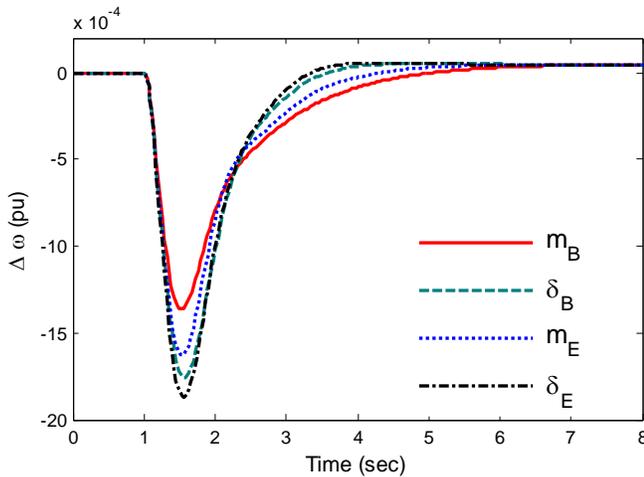
Fig. 15. Speed deviation response for Case-VI (step increase in Vref)

Case VII: Effect of parameter variation on the performance of UPFC-based damping controllers

In the design of damping controllers for any power system, it is extremely important to investigate the effect of variation of system parameters on the dynamic performance of the system. In order to examine the robustness of the damping controllers to variation in system parameters, a 25% decrease in machine inertia constant and 30% decrease of open circuit direct axis transient time constant is considered. The system response with the above parameter variations for a step increase in mechanical power is shown in Figs. 16-17 with all four alternative UPFC-based damping controllers.
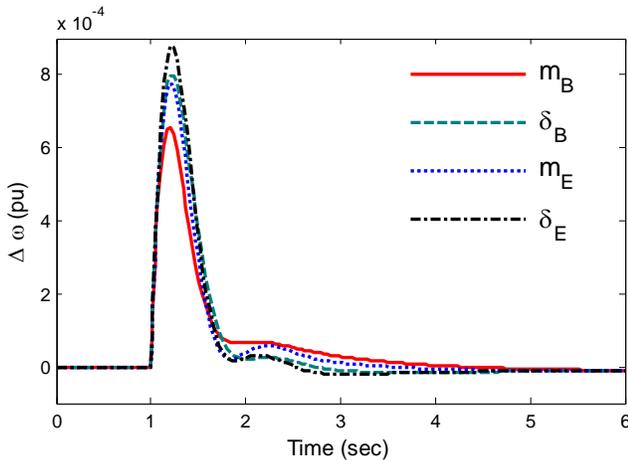
Fig. 16. Speed deviation response for Case-VII (25% decrease in M)
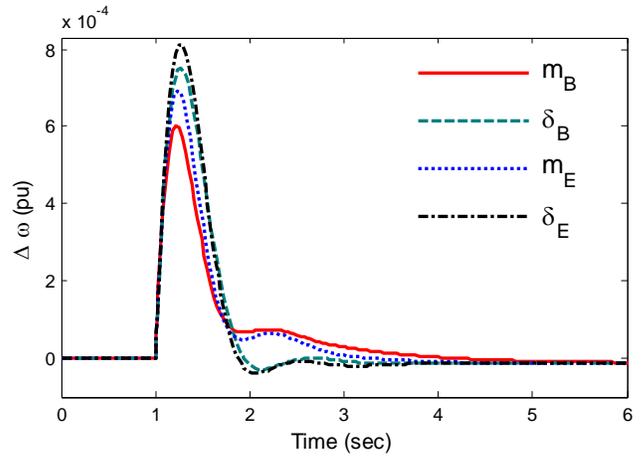
Fig. 17.Speed deviation response for Case-VII (30% decrease in T'$_{do}$)

It can be concluded from these Figs. that all four alternative damping controllers provide satisfactory damping performance with parameter variation. However, the performance of m_B - based UPFC damping controller seems to be slightly better among the four alternatives.

VI. CONCLUSION

In this study, a real-coded genetic algorithm optimization technique is employed for the design of UPFC-based damping controllers. The design problem is transferred into an optimization problem and RCGA is employed to search for the optimal UPFC-based controller parameters. The performance of the four alternatives UPFC based damping controller namely modulating index of series inverter (m_B), modulating index of shunt inverter (m_E), phase angle of series inverter (δ_B) and phase angle of the shunt inverter (δ_E) have been investigated under various disturbances and parameter variations. Simulation results are presented and compared with a conventional phase compensation technique for tuning the damping controller parameters to show the superiority of the proposed design approach. Investigations show that the damping control by changing the modulating index of series inverter (m_B), provide slightly better performance among the four alternatives.

APPENDIX
Static System data: All data are in pu unless specified otherwise.

Generator: M = 8.0 s., D = 0, $X_d = 1.0$, $X_q = 0.6$,

$$X'_d = 0.3, T'_{do} = 5.044, P_e = 0.8, V_t = V_b = 1.0$$

Excitor: $K_A = 100$, $T_A = 0.01$ s

Transformer: $X_{tE} = 0.1$, $X_E = X_B = 0.1$

Transmission line: $X_{Bv} = 0.3$, $X_e = 0.5$

UPFC parameters: $m_E = 0.4013$, $m_B = 0.0789$,

$$\delta_E = -85.3478^0, \delta_B = -78.2174^0, V_{DC}=2.0, C_{DC}=1.0$$

REFERENCES

[1] P. Kundur, Power System Stability and Control. New York: McGraw-Hill, 1994.

[2] N. G. Hingorani, L. Gyugyi, Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems, IEEE Press, New York, 2000.

[3] Y. H Song, T. A. Johns, Flexible AC Transmission Systems (FACTS), IEE, London, 2000.

[4] R. M Mathur, R. K. Verma, Thyristor-based FACTS Controllers for Electrical Transmission Systems, IEEE Press, Piscataway, 2002.

[5] L Gyugyi. 'Unified Power Flow Control Concept for Flexible ac Transmission Systems.' IEE Proceedings-C, Vol. 139, No 4, pp. 323-331, 1992.

[6] L. Gyugyi et. al., "The Unified Power Flow Controller: A New Approach to Power Transmission Control", IEEE Trans. Power Delv. Vol. 10, No. 2, pp. 1085-1093, 1995.

[7] Sidhartha Panda, "Application of non-dominated sorting genetic algorithm-II technique for optimal FACTS-based controller design", Journal of the Franklin Institute, doi:10.1016/j.jfranklin.2010.03.013.

[8] Sidhartha Panda, S.C. Swain, P.K. Rautray, R. Mallik, G. Panda, "Design and analysis of SSSC-based supplementary damping controller", Simulation Modelling Practice and Theory, doi: 10.1016/j.simpat.2010.04.007.

[9] Sidhartha Panda, "Multi-Objective Evolutionary Algorithm for SSSC-Based Controller Design", Electric Power System Research, Vol. 79, Issue 6, pp. 937-944, 2009.

[10] Sidhartha Panda, "Differential Evolutionary Algorithm for TCSC-Based Controller Design", Simulation Modelling Practice and Theory, Vol. 17, pp. 1618-1634, 2009.

[11] Sidhartha Panda, "Multi-objective Non-Dominated Shorting Genetic Algorithm-II for Excitation and TCSC-Based Controller Design", Journal of Electrical Engineering, Vol. 60, No. 2, pp. 87-94, 2009.

[12] Sidhartha Panda, N.P. Padhy, "Comparison of Particle Swarm Optimization And Genetic Algorithm for FACTS-Based Controller Design", Applied Soft Computing, Vol. 8, Issue 4, pp. 1418-1427, 2008.

[13] M. Nambu and Y. Ohsawa, "Development of an advanced power system stabilizer using a strict linearization approach", IEEE Transactions on Power Systems, Vol. 11, pp. 813-818, 1996.

[14] E. V. Larsen and D. A. Swann "Applying Power System Stabilizers Part II: Performance Objectives and Tuning Concepts', IEEE Transactions on Power Apparatus and Systems, Vol. 100, pp. 3025-3033, 1981.

[15] N. Tambey, M.L Kothari, "Unified Power Flow Controller (UPFC) Based Damping Controllers for Damping Low Frequency Oscillations in a Power System", IE(I) Journal-EL, Vol. 84, pp. 35-41, 2003.

[16] M. E. About-Ela, A. A. Sallam, J. D. McCalley and A. A. Fouad, "Damping controller design for power system oscillations using global signals', IEEE Transactions on Power Systems, Vol. 11, pp. 767-773, 1996.

[17] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989.

[18] Sidhartha Panda, A.K. Baliarsingh, R.K. Sahu, "Multi-objective Optimization Technique for TCSC-Based Supplementary Damping Controller Design", Proceedings World Congress on Nature & Biologically Inspired Computing (NaBIC 2009), 21st -23rd Dec., 2009 Bhubaneswar, India., pp. 1065-1070, 2009.

[19] Sidhartha Panda, C. Ardil, "Real-Coded Genetic Algorithm for Robust Power System Stabilizer Design", International Journal of Electrical, Computers and System Engineering, Vol. 2, No. 1, pp. 6-14, 2008.

[20] Sidhartha Panda, S.C. Swain, A.K. Baliarsingh, "Real-Coded Genetic Algorithm for Robust Coordinated Design Of Excitation And SSSC-Based Controller", Journal of Electrical Engineering, Vol. 8, Issue 4, pp. 31-38, 2008.

[21] S. Panda, S.C. Swain, A.K. Baliarsingh, C. Ardil, "Optimal Supplementary Damping Controller Design For TCSC Employing RCGA", International Journal of Computational Intelligence, Vol. 5, No. 1, pp. 36-45, 2009.

[22] H F Wang. 'A Unified Model for the Analysis of FACTS Devices in Damping Power System Oscillations — Part III : Unified Power Flow Controller.' IEEETransactions on Power Delivery, Vol. 15, no 3, July 2000, p 978.

[23] N. Tambey, M.L Kothari, "Damping of Power System Oscillations with Unified Power Flow Controller (UPFC)", IEE Proc. Gener. Trans. Distrib., Vol. 150, No. 2, pp. 129-140, 2003.

Data Locating In Real Time Cloud Based Service & Market Oriented Architectures

Boopathy.D
Research Scholar
Dr.G.R.D College of Science
Coimbatore
Tamilnadu

R.Srividhya
Assistant Professor
Dr.G.R.D College of Science
Coimbatore
Tamilnadu

Abstract

Cloud computing technology is a new concept of providing dramatically scalable and virtualized resources. It implies a service oriented architecture, reduced information technology overhead for the end-user, great flexibility, reduced total cost of ownership, on-demand services and many other things. When the Software as a Service Provider or Infrastructure as a Service provider may bankrupt or suddenly disappeared from the competitive market means, who will take responsible for customer data? One of the main concerns of customers is Cloud security and the threat of the unknown. The lack of physical access to servers constitutes a completely new and disruptive challenge for investigators. This paper represents the physical storage of data in end-user point of view with possible way and satisfies the standards and legal issues.

Keywords: Cloud computing, Security, Legal Aspects, Privacy access, Data Location.

1. Introduction:

Cloud computing is a natural evolution of the widespread adoption of virtualization, service-oriented architecture, autonomic, and utility computing. Details are abstracted from end-users, who no longer have need for expertise in, or control over, the technology infrastructure "in the cloud" that supports them. The NIST defines,"Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

Cloud computing models are of two types: Deployment model and Service model.

Deployment model is further classified into 4 type's namely private cloud, community cloud, public cloud and hybrid cloud.

In Private cloud, the cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise. In Community cloud, the cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise. In public cloud, the cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services. In hybrid cloud, the cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

Service model is also classified into three namely SaaS, PaaS, and IaaS.

1.1. Software as a Service (SaaS):

The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

1.2. Platform as a Service (PaaS):

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations

1.3. Infrastructure as a Service (IaaS):

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can

include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems; storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

2. Literature Review:

One of the most common compliance issues facing an organization is data location [1]. Use of an in-house computing center allows an organization to structure its computing environment and know in detail where data is stored and the safeguards used to protect the data. In contrast, a characteristic of many cloud computing services is that the detailed information of the location of an organization's data is unavailable or not disclosed to the service subscriber. This situation makes it difficult to ascertain whether sufficient safeguards are in place and whether legal and regulatory compliance requirements are being met. External audits and security certifications can, to some extent, alleviate this issue, but they are not a panacea. Once information crosses a national border, it is extremely difficult to guarantee protection under foreign laws and regulations [6].From the technical point of view, this evidence data can be available in three different states: at rest, in motion or in execution. Data at rest is represented by allocated disk space. Whether the data is stored in a database or in a specific file format, it allocates disk space. Furthermore, if a _le is deleted, the disk space is de-allocated for the operating system but the data is still accessible since the disk space has not been re-allocated and overwritten. This fact is often exploited by investigators which explore these de-allocated disk space on hard-disks. In case the data is in motion, data is transferred from one entity to another e.g. a typical file transfer over a network can be seen as a data in motion scenario. Several encapsulated protocols contain the data each leaving specific traces on systems and network devices which can in return be used by investigators. Data can be loaded into memory and executed as a process. In this case, the data is neither at rest nor in motion but in execution [10].Depending on the Cloud offer used, virtual IaaS instances do not have any persistent storage. In current Cloud environments CSP do not offer any verification process providing the ability for the customer to verify that the sensitive data stored on a virtual machine has been deleted exhaustively [10]. In the SaaS model, the enterprise data is stored outside the enterprise boundary, at the SaaS vendor end. Consequently, the SaaS vendor must adopt additional security checks to ensure data security and prevent breaches due to security vulnerabilities in the application or through malicious employees. This involves the use of strong encryption techniques for data security and fine-grained author- ization to control access to data [11]. Cloud computing moves the application software and

databases to the large datacenters, where the management of the data and services are not trustworthy. This unique attribute, however, poses many new security challenges (Cong Wang et al., 2009) [12]. Analysts' estimate that within the next five years, the global market for cloud computing will grow to \$95 billion and that 12% of the worldwide software market will move to the cloud in that period. To realize this tremendous potential, business must address the privacy questions raised by this new computing model (BNA, 2009) [13].

Yet, guaranteeing the security of corporate data in the ''cloud'' is difficult, if not impossible, as they provide different services like SaaS, PaaS, and IaaS. Each service has its own security issues (Kandukuri et al., 2009) [14]. Due to compliance and data privacy laws in various countries, locality of data is of utmost importance in much enterprise architecture (Softlayer, 2009) [15]. The security policies may entitle some considerations wherein some of the employees are not given access to certain amount of data. These security policies must be adhered by the cloud to avoid intrusion of data by unauthorized users (Blaze et al., 1999; Kormann and Rubin, 2000; Bowers et al., 2008) [16]. Data security is a significant task, with a lot of complexity. Methods of data protection, such as redaction, truncations, obfuscation, and others, should be viewed with great concern [17]. Data Loss/Leakage. Be it by deletion without a backup, by loss of the encoding key or by unauthorized access, data is always in danger of being lost or stolen [18]. In general, cloud users are not aware of the exact location of the datacenter and also they do not have any control over the physical access mechanisms to that data. Most well-known cloud service providers have datacenters around the globe. Some service providers also take advantage of their global datacenters. In cloud environment, data can be assigned a cost by the users based on the criticality of the data [19]. since a customer will not know where her data will be stored, it is important that the Cloud provider commit to storing and processing data in specific jurisdictions and to obey local privacy requirements on behalf of the customer; one needs to ensure that one customer's data is fully segregated from another customer's data; it is important that the Cloud provider has an efficient replication and recovery mechanism to restore data if a disaster occurs [20].

3. Problem Description:

Software as a service is mingled with Platform as a Service and Infrastructure as a Service.

The above diagram shows the Software as a Service providing process in simple style. The software as a service users got their account by free or for trial period or by payment mode. The mode of acquiring service may be different but the users can access their account and saved their confidential data into their account. In the

above figure, two concerns provide services and store their data in third parties. Who provides the IaaS?
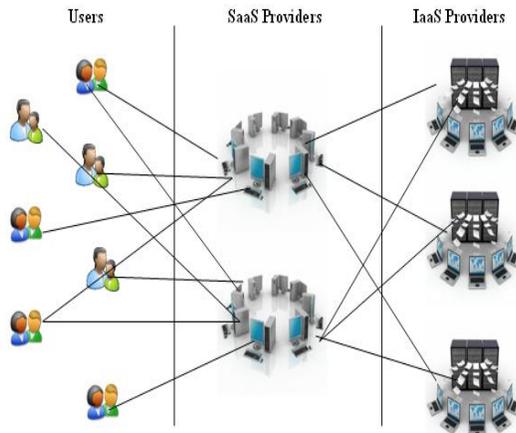
Fig: 1 SaaS providers Current Model

Here that IaaS person is unknown to the user. What is third party? Third party means, the users doesn't know where their data are stored and from where they access their stored data. The SaaS providers use many third parties to store their data because in online storage and accessing, the performance and security is important. For that performance issues they used more IaaS providers. So if more than one person into the service means, that company will name them as third party or fourth party? The users think their data are safe in online storage but the users don't know the status of their data. Because the data storage destination was not shown by the SaaS providers to the users. And the SLA is not satisfied the rules and regulations. Everything is fine now but in future when something happens (i.e.) when SaaS or IaaS may bankrupt or vanish from the service providing market, which time users may need their data back but they lose their data in reality. This paper presents some issues in Service Oriented Architecture with some solution model.

The data may contain anything but it is important to the users and its holders. The cloud computing handle the data in safe method but the assurance of the data quality and availability is very big question mark. The users must have to know the place of the data where it is saved and its details. Because at the time of downtime the SAAS provider said the problem is from hardware vendor part. But the actual problem may be in SAAS provider part. For the data availability and data assurance the "CLOUD DATA TAGG (CDT)" tool will help.

4. Cloud Data Tagg:
4.1 What is Cloud Data Tagg?

Cloud Data Tagg is process of allocating the tagg address to the data's which is handled in Cloud Computing areas. In reality the data's stored in cloud computing is strange to understand. Due to maintain the consistency and avoid the non-availability of data, the service providers store the data into multiple servers. Each and every server is interconnected and the data's were synchronized for future assistance, according to the country the provisional acts are differed from one to another. The reason for the difference, there is no common act which covers the whole word "Cloud Computing". The servers are installed in different places or the service provider may give contract to the IAAS service vendor that he has not servers in all over the world. So many IAAS service vendors were placed around the world to provide the full service. In this time data's were placed in many servers. The problem of non-availability is considered and raised only at the time issue. But at that time every thing will be exceed the limit of recovery or it result in data loss. Just think if the SAAS service provider or the IAAS provider became bankrupt means what happens to the data's stored in the server? If the clients know the places of data's at the service time itself means some problems will be solved from beginning itself.

4.1.1 Cloud Data Tagg (CDT) concept will act as a bridge between the user and the service both providers (both IAAS & SAAS). This is like a rendering service for the person who needs it. To basic service providers are two types

- IAAS – Clients – Users

- IAAS – SAAS – Clients – Users

The above types are followed and being followed in service providing part. The users will be the end users, because he/she is person who is going to utilize it and enjoy the full benefit of service. The CDT concept is followed means non-transparency is eradicated in data handling side. The IAAS and SAAS provider may accept mean it is possible to reach the prescribed goal. Why it need IAAS and SAAS support? The SAAS provider provides the software level service and the IAAS provider provides the hardware level service and he is person who stored the data's. If the both person are accept the concept means it will merged in their areas and start work.

So this tool will help the clients to track the data and knows the status of it.

Fig: 2 Cloud Data Tagg Module

4.2. CLOUD DATA TAGG (CDT) Modules:

Ξ   Data Tagg Allocator (DTA)

Ξ   Data Tagg LOCATOR (DTL)

Ξ   Data Tagg Tracker (DTT)

Ξ   Data Tagg Responser (DTR)

By using the above modules the clients can handle their data's easily.

4.2.1 DATA TAGG ALLOCATOR (DTA):

The DTA is one time process of allocating the serial code to the main data. It means the data is used in some specified software only so it sequential serial key is prepared one time and allocated it to the specified software (that is data accessing key). The user can select the auto saving method by their choices, the auto changes may done in following categories.

➢   Some specified data size reached means (Ex: Every 10 MB data will stored)

➢   Some specified time limit reached means (Ex: Every five mints data will stored)

➢   At the time of sign out or log out from the software the data stored.

While in above conditions auto save will done and each and every time of data saving it generate sequential numbers and tagged it with the saved data's and the tagg number will be
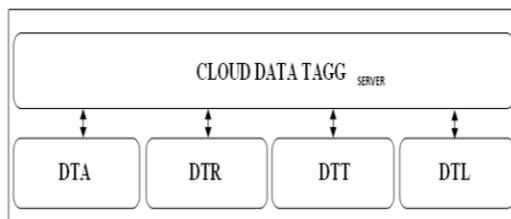
updated to the DTT. The DTA work is allocation of sequential data to saved data's and the sequential serial code will upload to the general data tagg server for future assistance. The each and every module access the sequential serial code from the server and it upload the details to the server. The details are saved and provide to the clients at their demand. The DTA process is accessed by SAAS provider only. The person who rendering service may able to make everything into transparency and also he can turn something into non-transparency, in such case the SAAS provider's SLA's are in non-

transparency method, due to some issues with IAAS provider. After they enter into the path of Cloud Data Tagg means the storage data's issues will came into transparency side.

At first time of allocating the space itself the software's also located in server. So that time itself the database became ready to store the data. So the software is already merged with the database. So for that database we provide the first sequential special code it used to find that database easily (code may contain 10 to 16 digits). After that the sub sequential codes will be randomly generated at the time of data saving and automatic data saving time. Time of auto saving will be recommended by the service provider and allocated by the user. (It may change in future by user only). In sub sequential code it tagg the data with date, day and time of saving will be noted. The DTA is only allocated the Data Tagg code, after that it will forward to the Data Tagg Database.

4.2.2 DATA TAGG LOCATOR (DTL):

The DTL is used to find the location of current saving data details by using its sequential serial code. So the user can know the location of data and also know the back up places of the data easily in sitting front of the computer. The DTL is connected with the Data Tagg Database for acquire the code details, the connection and method of accessing the code may very secure with the help of sequential code. The DTL is the current data locating module it is working in the CDT server by using the distributed computing method.

4.2.3 DATA TAGG TRACKER (DTT):

The DTT is another one important module. The DTT is simply known as history place of the specified data. It means from the first time a data is saved means the sequential serial number is added to this DTT module. From that time onwards that saved sequential data comes under tracking part. It simply surveillances the data and it recorded the information into the sequential file. It includes records of

➢   First time saved details

➢   The sub-data which are related to this data (includes the last saved data's)

➢   The accessing history of the data

➢   The data back up details

➢   The last retrieval details of the data (it shows the location of retrieval)

The data retrieval averages, counts and maximum hit of specified data's all are shown in this module. The DTT is connected with the Data Tagg

Database for acquire the code details, the connection and method of accessing the code may very secure.

4.2.4 DATA TAGG RESPONDER (DTR):

DTR is created for the purpose of alert the clients. The clients are not able to find out the downtime and also the data unavailability. So the DTR sends some alerts to the clients regarding the data server. The DTR is the module which responds the user's requirements. At the time of D&SLA signing it ask the users to select the set of services for the purpose of keep on touch with the data. That user's selected services will follow by the DTR and sends to users automatically. The main concept of DTR is response of the data while it is covers the limits or not.

- Data Tagg Ranking

- Data Tagg High-availability

- Data Tagg Non-availability

- Data Tagg Alerts and so on.

The DTR is connected with the Data Tagg Database for acquire the code details, the connection and method of accessing the code may very secure.

4.3. Data Tagg Database (DTD):

The Data Tagg Database (DTD) is the place where the sequential special code and sequential sub-code details may store in this database. The DTD accessed by the DTA for store the purpose of store sequential special code and sequential sub-code. The DTD accessed by the DTL for find out the location of the all required data's. The DTD accessed by the DTT for tracing the data's activities and this is the only module which keeps on connection with the DTD for future assistance. The DTD accessed by the DTR to respond the users selected requests on data's based service.

The DTD will be the centralized database which contains the full details of CDT server based service with distributed computing system.

4.5. Deployment & Service Level Agreement (D&SLA):

The Cloud Data Tagg may provide D&SLA to SAAS provider, why CDT need D&SLA means the module is going to provide the information's of stored data's as a service. The data's may contains anything which is related personnel, working area, or any other important thing. The clients are liable for the data what they stored. The CDT service is tracking the data from the beginning to ending, so the permission is needed from clients/users for track their data's and SAAS provider must give the permission to deploy this module into their software. And also SAAS provider must get the permission form IAAS provider to surveillances their storage data's which related to the SAAS provider privileges.

The Deployment Level covers and related to the SAAS and IAAS provider. Without their knowledge and support the module can't work successfully. The servers must show in the providing list with its transparency (otherwise it will make as transparency). The module deployment is made in SAAS and IAAS area. So, it related to the service vendors or service providers. The SLA is getting signed from the end-users of this service. The end-users are the person who utilizes the service from SAAS in the mode of software and the in-direct service from IAAS in the mode of storage. But he must comes under the SLA due to cross some issues for permit the service provider to track the data's and some other data's protected issues. The user's must clear in one thing, which is the service providers may use their data for their use it may include any reason. But in this place of CDT service, data is a data it contains anything which is related to the users or not. That data will be tagged, located, tracked, responded and the data stored details will be stored in database and give that tagg details to users without any fail and in right time.

4.6. CDT Service Model:
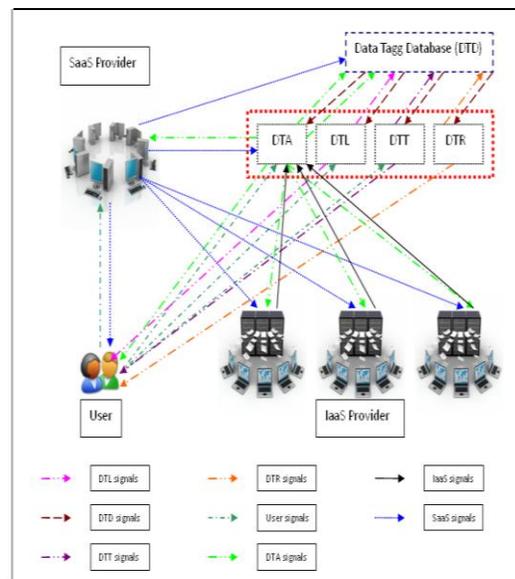
Fig: 3 Cloud Data Tagg Service Model

The above diagram explains CDT service. When the user request to create an account in SaaS provider then SaaS provider send that request to DTA for verification purpose. After that it sends the request to allocate the space for that user and that space details send to the DTA. Again DTA verifies the location of data storage and security categories. Later it enters those details to the

DTD. Now it sends the final statement to the SaaS provider. After that only the user able to get the account allocation information with the sequential code. When the user requests to locate the data he/she sends the request to the DTL, it provide that details to him. In the same way it is followed in the DTT system. And DTR send the automatic generated alerts regarding his/her accounts.

4.7. SEQUENTIAL DIAGRAM OF CDT SERVICE MODEL:
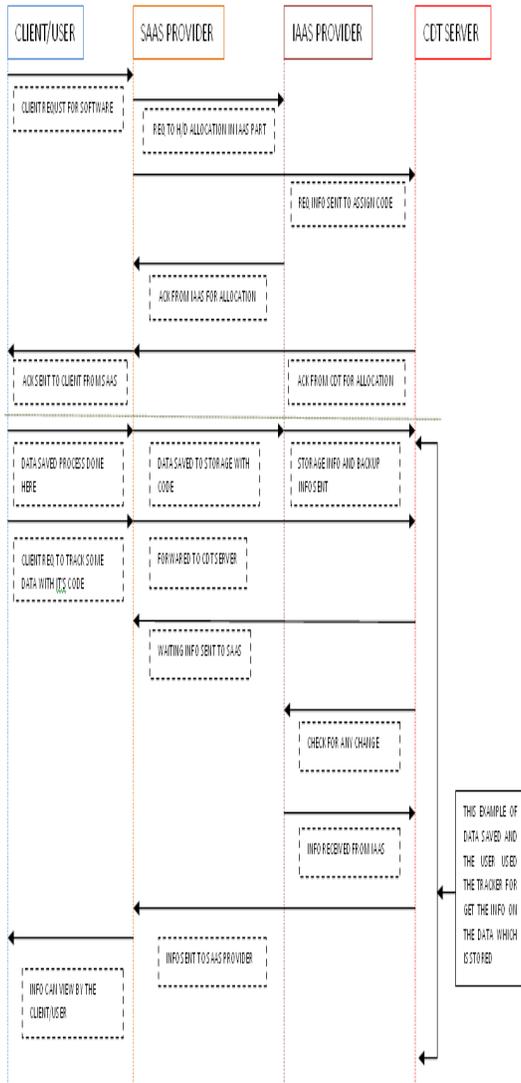
Fig: 4 CDT service model in sequential diagram

5. Future Work:

We are going to simulate the paper and we are attaching this model in all field based scenarios to verify the performance and security level. After that it will implemented in real world with high performance level and along with high level security. Still some modules are

under updating. Soon it will get updated and placed to check the working performance.

6. Simulation Results:
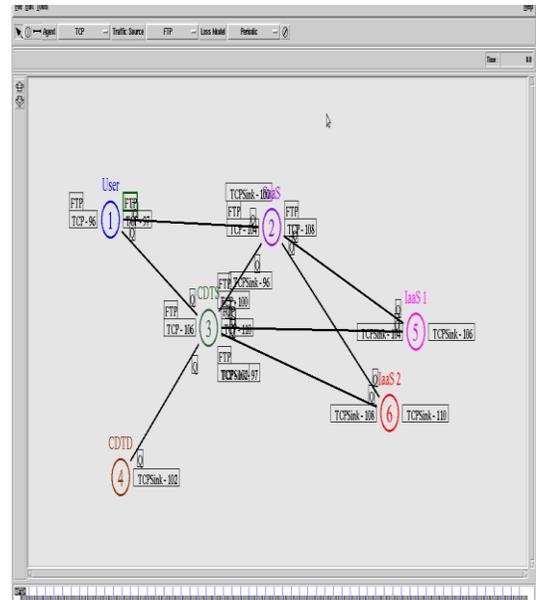
Fig: 5 DTA Design
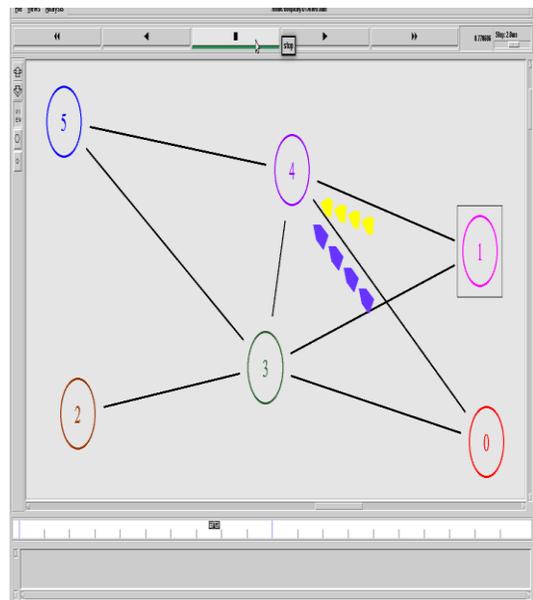
Fig: 6 DTA Output
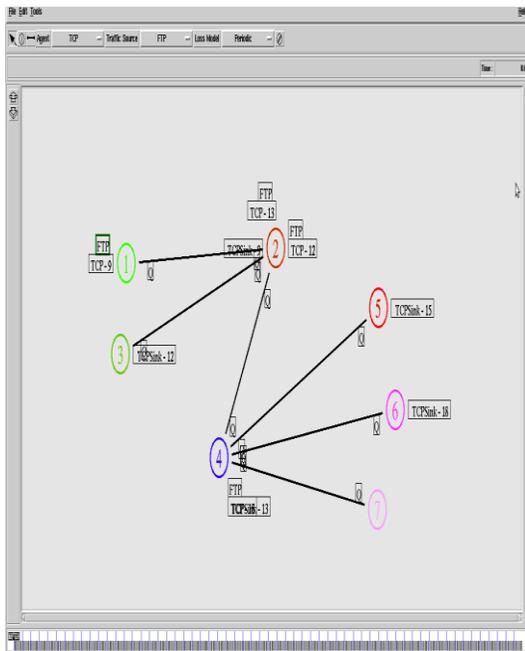
Fig: 7 DTL Design
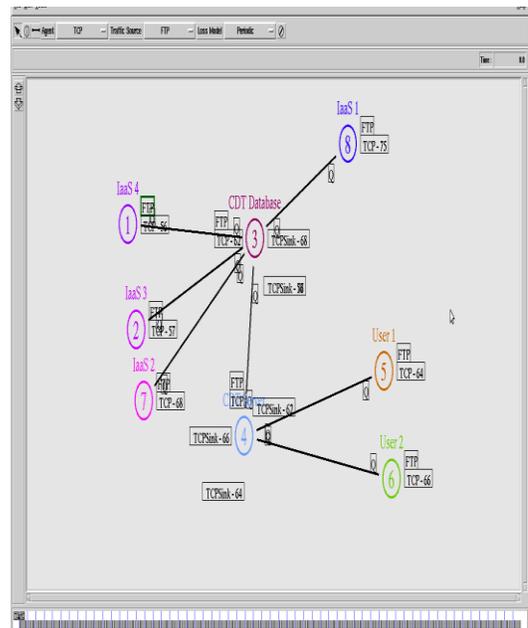
Fig: 9 DTT Desigh
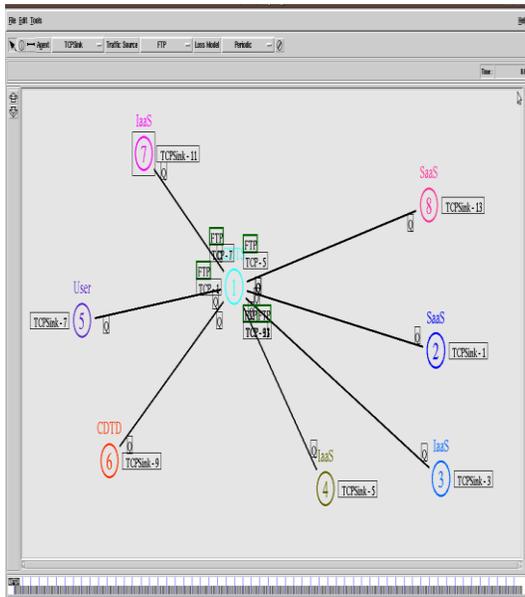
Fig: 8 DTR Design
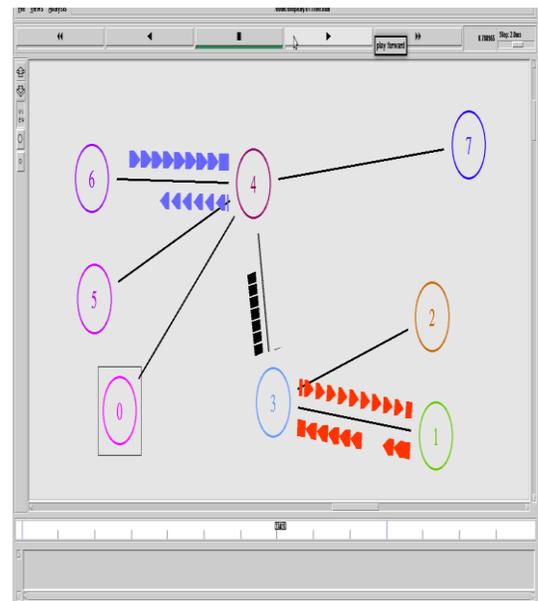
Fig: 10 DTT Output

7. Conclusion:

This paper provide a Service Oriented Architecture based model, it will give importance to security and performance, which is required by users of the cloud computing. This model is basic one for the other upcoming models of SOA, not only this model going too verified in Market Oriented Architecture (MOA). Soon this model will attach in SOA and MOA to verify it level of service.

References:

[1] DANISH JAMIL and HASSAN ZAKI, "SECURITY ISSUES IN CLOUD COMPUTING AND COUNTERMEASURES", IJEST, Vol. 3 No. 4 April 2011, and pg no: 2672-2676.

[2] Hyukho Kim, Hana Lee, Woongsup Kim, Yangwoo Kim, "A Trust Evaluation Model for QoS Guarantee in Cloud Systems", International Journal of Grid and Distributed Computing, Vol.3, No.1, March, 2010.

[3] Ghalem Belalem, Samah Bouamama, Larbi Sekhri, "An Effective Economic Management of Resources in Cloud Computing", JOURNAL OF COMPUTERS, VOL. 6, NO. 3, MARCH 2011, pg no: 404-411.

[4] Aishwarya C.S. and Revathy.S, "Insight into Cloud Security issues", UACEE International journal of Computer Science and its Applications, pg no: 30-33.

[5] Amreen Khan and KamalKant Ahirwar, "MOBILE CLOUD COMPUTING AS A FUTURE OF MOBILE MULTIMEDIA DATABASE", International Journal of ComputerScience and Communication, Vol. 2, No. 1, January-June 2011, pp. 219-221.

[6] Richard Chow, Philippe Golle, Markus Jakobsson, Ryusuke Masuoka, Jesus Molina Elaine Shi, Jessica Staddon, "Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control", CCSW'09, November 13, 2009.

[7] Muzafar Ahmad Bhat, Razeef Mohd Shah, Bashir Ahmad, "Cloud Computing: A solution to Geographical Information Systems (GIS)", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 2 Feb 2011, pg no: 594-600.

[8] Pardeep Kumar, Vivek Kumar Sehgal , Durg Singh Chauhan, P. K. Gupta and Manoj Diwakar, "Effective Ways of Secure, Private and Trusted Cloud Computing", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, May 2011, pg no: .412-421.

[9] Mladen A. Vouk, "Cloud Computing – Issues, Research and Implementations", Journal of Computing and Information Technology - CIT 16, 2008, 4, 235–246.

[10] Dominik Birk, "Technical Challenges of Forensic Investigations in Cloud Computing Environments" , January 12, 2011.

[11] A survey on security issues in service delivery models of cloud computing S. Subashini, V.Kavitha, Journal of Network and Computer Applications, 11 July 2010

[12] Toward publicly auditable secure cloud data storage services, Cong Wang Kui Ren Wenjing Lou Jin Li, Journal IEEE Network: The Magazine of Global Internetworking archive Volume 24 Issue 4, July-August 2010

[13] BNA. Privacy & security law report, 8 PVLR 10, 03/09/2009.Copyrigh 2009 by The Bureau of National Affairs, Inc.(800-372-1033), 2009 /http://www.bna.comS [accessed on:2November2009].

[14] Kandukuri BR, Paturi VR, Rakshit A. Cloud security issues. In: IEEE international conference on services computing, 2009, p. 517–20.

[15] Softlayer.Service Level Agreement and Master Service Agreement, 2009 /http:// www.softlayer.com/sla.html [accessed on:11December2009].

[16] Bowers KD, Juels A, Oprea A. HAIL: a high-availability and integrity layer for cloud storage, Cryptology ePrint Archive, Report 2008/489, 2008 /http://eprint.iacr. org [accessed on:18October2009].

[17] Survey on Cloud Computing Security, Shilpashree Srinivasamurthy, David Q. Liu, 2nd IEEE International Conference on *Cloud Computing,* 2010

[18] Security Guidance for Critical Areas of Focus in Cloud Computing, April 2009. DOI = http://www.cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf

[19] Towards Analyzing Data Security Risks in Cloud Computing Environments Amit Sangroya, Saurabh Kumar, Jaideep Dhok, and Vasudeva Varma, Springer-Verlag Berlin Heidelberg 2010

[20] Cloud Computing and Grid Computing 360-Degree Compared, Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu, Grid Computing Environments Workshop, 2008. GCE '08 2008

AUTHORS

Boopathy.D is Research Scholar of Computer Science Department, School of IT & Science, Dr. G R Damodaran College of Science, Coimbatore. He received his Master of Science degree in Information Technology in the year of 2010 from Bharathiar University, Coimbatore. His research area is Data warehousing and mining.

R.Srividhya is Assistant Professor of School of IT & Science, Dr.G R Damodaran College of Science, Coimbatore. She received her Master of Computer Application degree in the year of 2002 from Bharathiar University and she received her Master of Philosophy in the year of 2004 from Bharathiar University. She has worked in RVS College of Arts & Science for 3 years as Lecturer and currently working as Assistant Professor in School of IT & Science, Dr. G R Damodaran College of Science, Coimbatore. The research areas are Data warehousing and Mining, Wireless Networks and Digital Image Processing.

Applied lacunary interpolation for solving Boundary value problems

Faraidun K. Hamasalh
University of Sulaimani-School of Science Education-Department of Mathematics-Kurdistan-Iraq

Abstract

Interpolation by various types of splines is the standard procedure in many applications. In this paper shall discuss the function, two and fourth derivatives of spline interpolation as an alternative to polynomial spline interpolation at the all intervals. The method is appropriate and solving of initial and boundary value problems, the results revealing that method is very effective and accurate.

Keywords: Cauchy problem, Spline function, Initial and Boundary value problems, Taylor's expansion.

1. Introduction.

We consider the following initial and boundary values problem:

$$y''(x) = f(x, y(x), y'(x)) \, , \, x \in [0, 1], \; y(a) = y_1 \, , \, y'(a) = y_2'$$
$$y''(x) = f(x, y(x), y'(x)) \, , \, x \in [0, 1], \; y(a) = y_3 \, , \, y'(b) = y_4' \tag{1}$$

With the help of lacunary spline functions of type (0, 2, 4) see Faraidun (2010) [2], by using that $f \in C^{n-1}([0,1] \times R^2)$, $n \geq 2$ and that it satisfies the Lipschitz continuous

$$\left| f^{(q)}(x, y_1, y_1') - f^{(q)}(x, y_2, y_2') \right| \leq L \left\{ \left| y_1 - y_2 \right| + \left| y_1' - y_2' \right| \right\}, q=0,1,\ldots, n\text{-}1 \tag{2}$$

Also boundary value problems are satisfied, and for all $x \in [0,1]$ and for all real y_1, y_2, y_1', y_2'. These conditions ensure the existence of unique solution of the problem (1).

In [2] authors investigated the model (0, 2, 4) approximation by polynomial splines on box partitions in all intervals. The main computational advantage of this technique is its simple applicability for solving boundary value problems. We develop a new spline approximation method for solving the boundary value problems over the interval [a, b].

In section 2, we give a brief description of the method. The derivation of the difference schemes spline function has been given in Section 3, and also, we have shown the second-order accuracy method and convergence analysis are studied. We have solved two numerical examples to demonstrate the applicability of the methods in section 4. In the last section, the discussion on the results is given in Section 5.

2. Construct of approximate values:

Let $\quad \omega(h, \, y^{(r)}) = \underset{|x-x_1| \leq h}{Max} \left\{ \left| y^{(r)}(x) - y^{(r)}(x) \right| \right\} \quad , \quad r = 0,1, \ldots, 6.$ And let $\quad \overline{Y}_k^{(q)} : \quad \overline{y}_0^{(q)}, \overline{y}_1^{(q)}, \overline{y}_2^{(q)}, \ldots, \overline{y}_n^{(q)} ;$ $q = 0,1, \ldots, 6,$ be approximate to the exact values $Y_k^{(q)} : y_0^{(q)}, y_1^{(q)}, y_2^{(q)}, \ldots, y_n^{(q)} ; q = 0,1, \ldots, 6.$

Now from these approximate values we construct a spline function $\overline{S}_\Delta(x)$ wich interpolates to the set $\overline{Y}$ on the mesh Δ and approximate the solution y(x) of equation (1) as [4, 5]. The set $\overline{Y}^{(q)}$ is defined as:

$$\overline{y}_0 = y_0, \; \overline{y}'_0 = y'_0, \; \overline{y}_0^{(2+q)} = f^{(q)}(x_0, \, y_0 \, y'_0) \text{ where } q = 0 \; 1 \ldots, r \, .$$

$$\overline{y}_{k+1} = \overline{y}_k + h\overline{y}'_k + \int\limits_{x_k}^{x_{k+1}} \int\limits_{x_k}^{t} f[u, \, y_k^*(u), \, y_k^{**}{}'(u)] \, du \, dt \, ,$$

$$\overline{y}'_{k+1} = \overline{y}'_k + \int_{x_k}^{x_{k+1}} f[t, y_k^*(t), y_k^{**}{}'(t)]dt,$$

$$\overline{y}_{k+1}^{(q+2)} = f^{(q)}(x_{k+1}, \overline{y}_{k+1}, \overline{y}'_{k+1}), \quad q = 0,2,4., \quad k = 0, 1, 2, ..., m-1$$

and for $x_k \le x \le x_{k+1}$

$$y_k^*(x) = \sum_{j=0}^{r+2}(x-x_k)^j \frac{\overline{y}_k^{(j)}}{j!}, \qquad y_k^*{}'(x) = \sum_{j=0}^{r+1}(x-x_k)^j \frac{\overline{y}_k^{(j+1)}}{j!}$$

and $\qquad y'^{**}_{k+1}(x) = \overline{y}'_k + \int_{x_k}^{x_1} f[t, y_k^*(t), y_k^*{}'(t)]dt$

Using these approximate values $\overline{Y}_k^{(q)}$ $(q = 0,2,4., \ k = 0, 1, 2, ...,m)$ and $\overline{y}'_0, \overline{y}'_m$ on the bases of [2, 3], we construct the lacunary spline function $\overline{S}_\Delta(x)$ of the type $(0, 2, 4)$ ($\overline{S}_\Delta(x) = \overline{S}_k(x)$ if $x_k \le x \le x_{k+1}$) and denote by $\overline{S}_{n,6}^5$ the class of six degree splines $\overline{S}(x)$ as the following:

$$G(x) = \begin{cases} \overline{S}_\Delta(x_k) = \overline{y}_k \\ \overline{S}_\Delta^{(q)}(x_k) = \overline{y}_k^{(q)} \end{cases} \tag{3}$$

Where $q = 2, 4$ and $k = 0,1, 2, ...,m$, the existence and uniqueness of the above spline function have been shown in [2],

$$\overline{S}_0 = \overline{y}_0 + (x-x_0)\overline{y}'_0 + \frac{(x-x_0)^2}{2}\overline{y}_0'' + \overline{a}_{0,3}\overline{y}_o^{(3)} + \frac{(x-x_0)^4}{24}\overline{y}_o^{(4)} + (x-x_0)^5 \overline{a}_{0,5} + (x-x_0)^6 \overline{a}_{0,6} \tag{4}$$

Let us examine now intervals $[x_i, x_{i+1}]$, $i=1,2,..., n$-2., Defined $\overline{S}_i(x)$ as:

$$\overline{S}_i(x) = \overline{y}_i + (x-x_i)\overline{a}_{i,1} + \frac{(x-x_i)^2}{2}\overline{y}_i'' + (x-x_i)^3 \overline{a}_{i,3} + \frac{(x-x_i)^4}{24}\overline{y}_i^{(4)} + (x-x_i)^5 \overline{a}_{i,5} + (x-x_i)^6 \overline{a}_{i,6} \tag{5}$$

Here

$$\overline{a}_{0,3} = \frac{5h^{-3}}{3}(\overline{y}_1 - \overline{y}_0) - \frac{1}{18h}(2\overline{y}_1'' + 13\overline{y}_0'') - \frac{5}{3}h\ \overline{y}'_0 + \frac{h}{216}(\overline{y}_1^{(4)} - 4\overline{y}_0^{(4)});$$

$$\overline{a}_{0,5} = h^{-5}(\overline{y}_0 - \overline{y}_1) + \frac{h^{-3}}{6}(\overline{y}_1'' + 2\overline{y}_0'') + h^{-4}\overline{y}'_0 - \frac{h^{-1}}{360}(4\overline{y}_1^{(4)} + 11\overline{y}_0^{(4)});$$

and

$$\overline{a}_{0,6} = \frac{h^{-6}}{3}(\overline{y}_1 - \overline{y}_0) - \frac{h^{-4}}{18}(\overline{y}_1'' + 2\overline{y}_0'') - \frac{h^{-5}}{3}\overline{y}_0'' + \frac{h^{-2}}{1080}(7\overline{y}_1^{(4)} + 8\overline{y}_0^{(4)});$$

Also

$$\overline{a}_{i,1} + \overline{a}_{i+1,1} = 2h^{-1}(\overline{y}_{i+1} - \overline{y}_i) + \frac{h}{6}(\overline{y}_{i+1}'' - \overline{y}''_i) - \frac{h^3}{360}(\overline{y}_{i+1}^{(4)} - \overline{y}_i^{(4)});$$

$$\overline{a}_{i,3} = -\frac{5}{3}h^{-1}\overline{a}_{i,1} + \frac{5}{3}h^{-3}(\overline{y}_{i+1} - \overline{y}_i) - \frac{1}{18}h^{-2}(2\overline{y}_{i+1}'' + 13\overline{y}_i'') + \frac{h}{216}(\overline{y}_{i+1}^{(4)} - 4\overline{y}_i^{(4)});$$

$$\overline{a}_{i,5} = h^{-4}\overline{a}_{i+1,1} - h^{-5}(\overline{y}_{i+1} - \overline{y}_i) + \frac{h^{-3}}{6}(\overline{y}_{i+1}'' + 2\overline{y}_i'') - \frac{h^{-1}}{360}(4\overline{y}_{i+1}^{(4)} + 11\overline{y}_i^{(4)});$$

and

$$\overline{a}_{i,6} = -\frac{h^{-5}}{3}\overline{a}_{i,1} + \frac{h^{-6}}{3}(\overline{y}_{i+1} - \overline{y}_i) - \frac{h^{-4}}{18}(\overline{y}_{i+1}'' + 2\overline{y}_i'') + \frac{h^{-2}}{1080}(7\overline{y}_{i+1}^{(4)} + 8\overline{y}_i^{(4)}).$$

Similarly for the last interval $[x_n, x_{n-1}]$, we can define approximate values of $\overline{S}_n(x)$.

3. Convergence of a spline functions to a solution:

A key ingredient in the development of our estimates is the following theorem which gives a bound on the size of a polynomial on a spline function $\bar{S}_\Delta(x)$ in terms of its values on a discrete subset which is scattered in the values of y_k ($k = 0,1,2,....,m$) of a problem (1).

Theorem 1: Let $\bar{y}_k^{(q)}$ ($q = 0, 2, 4; k = 0, 1, 2..., m$) be the approximate values defined above. Then the following estimates of spline function $\bar{S}_\Delta(x)$ are valid:

(i) $\left| S_k^{(q)}(x) - \bar{S}_k^{(q)}(x) \right| \leq CI_q h^{6-q} \omega_6(h)$; for $q = 0,1,....,6$, $k = 0, 1,...,m-2$

where C_q denote the difference constants dependent of h.

(ii) $\left| y_k^{(q)}(x) - \bar{S}_k^{(q)}(x) \right| \leq H_q \sum_{j=0}^{q-1} \theta^{-j} (h_j + h^j \|D^j f\|_p)$; for $q = 0,1,....,6$, where $y(x)$ is a solution of problem (1) and D_q denote the difference constants dependent of h.

Proof: (i) From theorem 1 of [1] and equation (3), we have

$$S_0(x) - \bar{S}_0(x) = (x - x_0)^3 (a_{0,3} - \bar{a}_{0,3}) + (x - x_0)^5 (a_{0,5} - \bar{a}_{0,5}) + (x - x_0)^6 (a_{0,6} - \bar{a}_{0,6}) \tag{6}$$

Where

$$a_{0,3} - \bar{a}_{0,3} = \frac{5}{3h^3}(y_1 - \bar{y}_1) - \frac{1}{9h}[y_1'' - \bar{y}_1''] + \frac{h}{216}[y_1^{(4)} - \bar{y}_1^{(4)}]$$

implies that

$$\left| a_{0,3} - \bar{a}_{0,3} \right| \leq \frac{1}{216}(C_1 + 24C_2 + 72C_3)\,\omega_6(h) = \frac{1}{216} I_1 \,\omega_6(h)$$

where $I_1 = C_1 + 24C_2 + 72C_3$ and C_1, C_2 and C_3 are constants dependent of h.
Similarly

$$\left| a_{0,5} - \bar{a}_{0,5} \right| \leq \frac{1}{h^5}\left|y_1 - \bar{y}_1\right| + \frac{1}{6h^3}\left|y_1'' - \bar{y}_1''\right| + \frac{1}{90h}\left|y_1^{(4)} - \bar{y}_1^{(4)}\right|$$

$$\leq \frac{1}{90}(C_4 + 15C_5 + 90C_6)\omega_6(h) = \frac{1}{90} I_2 \omega_6(h)$$

where $I_2 = C_4 + 15C_5 + 90C_6$ and C_4, C_5 and C_6 are constants dependent of h.
and

$$\left| a_{0,6} - \bar{a}_{0,6} \right| \leq \frac{1}{3h^6}\left|y_1 - \bar{y}_1\right| + \frac{1}{18h^4}\left|y_1'' - \bar{y}_1''\right| + \frac{7}{1080h^2}\left|y_1^{(4)} - \bar{y}_1^{(4)}\right|$$

$$\leq \frac{1}{1080}(7C_7 + 60C_8 + 360C_9)\omega_6(h) = \frac{1}{1080} I_3 \omega_6(h)$$

where $I_3 = 7C_7 + 60C_8 + 360C_9$ and C_7, C_8 and C_9 are constants dependent of h.
And hence

$$\left| S_0(x) - \bar{S}_0(x) \right| \leq h^3 \left|a_{0,3} - \bar{a}_{0,3}\right| + h^5 \left|a_{0,5} - \bar{a}_{0,5}\right| + h^6 \left|a_{0,6} - \bar{a}_{0,6}\right|$$

$$\leq I \,\omega_6(h)$$

Where $I = I_1 + I_2 + I_3$, dependent of h.
By taking the first derivative of equation (5), we have

$$\left|S_0'(x) - \bar{S}_0'(x)\right| \le \frac{2}{h}\left|y_1 - \bar{y}_1\right| + \frac{5h}{6}\left|y_1'' - \bar{y}_1''\right| - \frac{h^3}{360}\left|y_1^{(4)} - \bar{y}_1^{(4)}\right|$$

$$\le \frac{1}{360}(\overline{C}_1 + 300\overline{C}_2 + 720\overline{C}_3)\,\omega_6(h) = \frac{1}{360}\bar{I}_1\omega_6(h)$$

and by successive differentiations obtain

$$\left|S_0^{(q)}(x) - \bar{S}_0^{(q)}(x)\right| \le I_q\,h^{6-q}\,\omega_6(h); \text{for } q = 0,1,\ldots,6.$$

This proves (i) for $k = 0$ and $x \in [x_0, x_1]$. Further more in the interval $[x_{k-1}, x_k]$

$$S_k(x) - \bar{S}_k(x) = (x - x_k)(a_{k,1} - \bar{a}_{k,1}) + (x - x_k)^3(a_{k,3} - \bar{a}_{k,3}) + (x - x_k)^5(a_{k,5} - \bar{a}_{k,5})^5$$

$$+ (x - x_k)^6(a_{k,6} - \bar{a}_{k,6})$$

From [2, 6], it's clear that, to show

$$a_{k,1} - \bar{a}_{k,1} = \frac{2}{h^2}(y_1 - \bar{y}_1) + \frac{h}{6}(y_1'' - \bar{y}_1'') + \frac{h^3}{360}[y_1^{(4)} - \bar{y}_1^{(4)}]$$

implies that

$$\left|a_{k,2} - \bar{a}_{k,2}\right| \le \frac{1}{360}(2C_0^* + 6C_1^* + C_2^*)\,\omega_6(h) = \frac{1}{360}I_1^*\,\omega_6(h)\ ;$$

where I_1^* and C_0^*, C_1^* and C_2^* be a constants dependent of h.

Similarly

$$\left|a_{k,3} - \bar{a}_{k,3}\right| \le \frac{5}{3h}\left|\bar{a}_{k,1} - a_{k,1}\right| + \frac{5}{3h^3}\left|y_{k+1} - \bar{y}_{k+1}\right| + \frac{1}{9h^2}\left|\bar{y}_{k+1}^{(2)} - y_{k+1}^{(2)}\right| + \frac{h}{216}\left|y_{k+1}^{(4)} - \bar{y}_{k+1}^{(4)}\right|$$

$$\le \frac{1}{216}(360C_3^* + 360C_4^* + 24C_5^* + C_6^*)\,\omega_6(h) = \frac{1}{216}I_2^*\,\omega_6(h)$$

where I_1^* and C_3^*, C_4^*, C_5^* and C_6^* be a constants dependent of h.

And also

$$\left|a_{k,5} - \bar{a}_{k,5}\right| \le I_3^*\,\omega_6(h)\ ;\ \left|a_{k,6} - \bar{a}_{k,6}\right| \le I_4^*\,\omega_6(h), \text{ where } I_2^* \text{ and } I_4^* \text{ are dependent of h.}$$

and by taking the successive differentiation, we obtain

$$\left|S_k^{(q)}(x) - \bar{S}_k^{(q)}(x)\right| \le I_q\,h^{6-q}\,\omega_6(h); \text{for } q = 0,1,\ldots,6. \text{ Which is prove (i) for } k = 0, 1,\ldots,m-2.$$

We can repeat the same manner in above for $k = m-1$.

Proof of theorem 1 (ii):

$$\left|y^{(q)}(x) - \bar{S}_\Delta^{(q)}(x)\right| \le C\left(\left\|y^{(q)}(x) - S_\Delta^{(q)}(x)\right\|_{L\infty} + \left\|S_\Delta^{(q)}(x) - \bar{S}_\Delta^{(q)}(x)\right\|_{L\infty}\right)$$

From theorem 2 [4], and after some derivations the following estimates are valid

$$\left\|y^{(q)}(x) - \bar{S}_\Delta^{(q)}(x)\right\|_{L\infty} \le C_q h^{6-q}\omega_6(h)\ , \text{ where } h = \theta\Delta c \text{ and } \omega_6(f;h)_p \le C\sum_{j=0}^{r-1}\theta^{-j}(h_r + h^r\|D^r f\|_p) \qquad (7)$$

Using equation (7) and estimate in (i), we have

$$\left|y^{(q)}(x) - \bar{S}_\Delta^{(q)}(x)\right| \le C_q h^{6-q}\omega_6(h) + I_q h^{6-q}\omega_6(f;h)$$

$$= (C_q + I_q)h^{6-q}\omega_6(f;h) = H_q h^{6-q}\omega_6(f;h) \le H_q\sum_{j=0}^{q-1}\theta^{-j}(h_j + h^j\|D^j f\|_p)$$

Which is proves (ii).

Theorem 2: If the function f in Cauchy's problem (1) satisfies conditions (2) and (3), then the following inequalities are hold:

$$\left\| \bar{S}_0''(x) - f[x, \bar{S}_0(x), \bar{S}_0'(x)] \right\|_{Lp} \le I_{0,2}^* \, \omega_6(h) \text{ where } I_{0,2}^* \text{ is constants dependent of h and } x \in [x_0, \, x_1],$$

$$\left\| \bar{S}_k''(x) - f[x, \bar{S}_k(x), \bar{S}_k'(x)] \right\|_{Lp} \le I_{k,2}^* \, \omega_6(h) \text{ where } I_{k,2}^* \text{ is constants dependent of h and } x \in [x_{k-1}, \, x_k],$$

$$\left\| \bar{S}_{m-1}''(x) - f[x, \bar{S}_{m-1}(x), \bar{S}_{m-1}'(x)] \right\|_{Lp} \le I_{m-1,2}^* \, \omega_6(h) \quad \text{where} \quad I_{m-1,2}^* \text{is constants dependent of h and}$$

$x \in [x_{m-1}, \, x_m]$.

Proof: Using condition (1), (2) and (3), we have

$$\left\| D^q(f(x) - y(x)) \right\|_{Lp} \le C_1 \omega_q(f; b-a) \text{ and } \left\| D^q y(x) \right\|_{Lp} \le C_2 \, \omega_q(f; 1)$$

by the Taylor expanssion of y about zero, then

$$\left| D^q(y(x) - \bar{S}_\Delta(x)) \right| \le \int_0^u \left| D^{q+1}(y - \bar{S}_\Delta)(u) \right| du \le \left\| D^{q+1}(y - \bar{S}_\Delta) \right\|_{Lp}$$

$$\le 2^{\frac{r}{p}} \left\| D^q y \right\|_{Lp} \le C_3 \, \omega_q(f; 1)$$

$$\left\| D^q(\bar{S}_0(x) - f(x)) \right\|_{Lp} \le \left\| D^q(\bar{S}_0 - y) \right\|_{Lp} + \left\| D^q y - f \right\|_{Lp} \le I_{0,2}^* \omega_6(f; 1) \text{ ; where } q = 2.$$

Similarly for each the intervals can be proving it.

 4. Numerical results:

   In this section, the method discussed in section 2 and 3 were tested on two problems, and the absolute errors in the analytical solution were calculated. Our results confirm the theoretical analysis of the methods with the initial and boundary value problems. For different starting points observed same convergence point with or less iterations, see [7].

Problem (1): we consider that the second order boundary value problem $y'' + y = 0$ where $x \in [0,1]$ and $y(0) = 1, \, y'(0) = 1$.

Problem (2): Let $y''' - y' = 2\cos(x)$ where $y(0) = 3, \, y'(0) = 2, \, y'(1) = 2$.

   It turns out that the six degree spline which presented in this paper, yield approximate solution that is $O(h^6)$ as stated in Theorem 1. The results are shown in the Table 1 and Table 2 for different step sizes h.

Table 1   Absolute maximum error for the derivatives $\bar{S}(x)$.

h	$\left\| \bar{s}'(x) - y'(x) \right\|_\infty$	$\left\| \bar{s}'''(x) - y'''(x) \right\|_\infty$	$\left\| \bar{s}^{(5)}(x) - y^{(5)}(x) \right\|_\infty$	$\left\| \bar{s}^{(6)}(x) - y^{(6)}(x) \right\|_\infty$
0.1	67.67×10^{-10}	26.06×10^{-7}	73×10^{-4}	11.33×10^{-2}
0.01	64.71×10^{-16}	22.41×10^{-11}	72.05×10^{-6}	11.1×10^{-3}
0.001	44.04×10^{-14}	22.2×10^{-7}	26.64×10^{0}	53.2×10^{3}

Table 2   Absolute maximum error for the derivatives $\bar{S}(x)$.

h	$\left\| \bar{s}'(x) - y'(x) \right\|_\infty$	$\left\| \bar{s}'''(x) - y'''(x) \right\|_\infty$	$\left\| \bar{s}^{(5)}(x) - y^{(5)}(x) \right\|_\infty$	$\left\| \bar{s}^{(6)}(x) - y^{(6)}(x) \right\|_\infty$
0.1	69×10^{-8}	25.32×10^{-5}	50.33×10^{-2}	82.005×10^{-1}
0.01	72.62×10^{-13}	25.38×10^{-8}	53.4×10^{-3}	87.01×10^{-1}
0.001	66.61×10^{-18}	22.2×10^{-11}	5×10^{-3}	79×10^{-1}

5. Conclusion:

An efficient and accurate numerical scheme based on the Interpolation method proposed for solving initial and boundary value problems. The Lacunary interpolation method was employed to reduce the problem to the solution of differential equations. Illustrative examples are presented in Table 1 and 2, were given to demonstrate the validity and applicability of the method with the less errors bounded.

References:

[1] Abbas Y. Al Bayati, Rostam K. Saeed and Faraidun K. Hama-Salh (2009), The Existence, Uniqueness and Error Bounds of Approximation Splines Interpolation for Solving Second-Order Initial Value Problems, Journal of Mathematics and Statistics   5 (2):123-129, ISSN 1549-3644.

[2] Faraidun K. Hama-Salh (2010), Numerical Solution for Fifth Order Initial Value Problems using Lacunary Interpolation, Acceptance letter for Publication from Dohuk Journal.

[3] Faraidun K. Hama-Salh, Karwan H. F. Jwamer (2011), Cauchy problem and Modified Lacunary Interpolations for Solving Initial Value Problems, Int. J. Open Problems Comp. Math. Vol. 4, No. 1, pp. 172-183.

[4] Gyovari, J. (1984), Cauchy problem and Modified Lacunary Spline functions, Constructive Theory of Functions, Vol.84, pp. 392-396.

[5] Karwan H. F. Jwamer (2005), Solution of Cauchy's problem by using spline interpolation, Journal of Al-Nahrain University, Vol. 8(2), December, pp.97-99.

[6] Saxena, A. (1987), Solution of Cauchy's problem by deficient lacunary spline interpolations, Studia Univ. BABES-BOLYAI MATHEMATICA, Vol.XXXII, No.2,   60-70.

[7] S.R.K. Lyengar and R.K. Jain, Numerical Method (2009); NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS, 4835/24, Ansari Road, Daryaganj, New Delhi - 110002 ISBN (13), 978-81-224-2707-3.

Design Methodology of Current Buffer based Two Stage CMOS Op-Amp with Compensation Strategy

Mr.Mayank Kumar Rai[1], Mr.Gaurav Mitra[2], Mr.Anubhav Kumar Tiwari[3]

[1] Thapar University, Patiala, India
[2] Bharati Vidyapeeth's College of Engg., Delhi, India
[3] Shobhit University, Meerut, India

Abstract

High Bandwidth Operational Amplifiers are needed for many applications. The Design methodology with current buffer overcomes the drawbacks in design strategies of nulling resistor and voltage buffer. The approach here provides improved gain bandwidth product and a gain of 42 dB on 0.5μm technology when operated on a supply voltage of 2.5 volts.

Keywords—CMOS analog integrated circuits, Common Source Stage, Compensation Capacitor, Op-Amp.

Introduction

Operational amplifiers are amplifiers (controlled sources) that have sufficiently high forward gain so that when negative feedback is applied, the closed-loop transfer function is practically independent of the gain of the op-amp. Most of the amplifiers do not have a large enough gain. Consequently, most CMOS op-amps use two or more gain stages [1]-[3].

The goal of compensation is to maintain stability when negative feedback is applied around the op-amp.

I. OP-AMP GAIN

Figure below shows block diagram of a two stage op-amp.
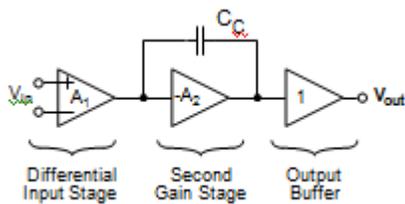
Fig 1 A Two stage Op-amp block Diagram.

First stage differential-to-single ended gain is given by

$$A_{v1} = g_{m1}(r_{ds2} \parallel r_{ds4}) \qquad (1)$$

where
Second stage gain is given by

$$A_{v2} = -g_{m7}(r_{ds6} \parallel r_{ds7}) \qquad (2)$$

Third stage is a source-follower and is only included if resistive loads need to be driven. If the load is purely capacitive in the case of integrated op-amps this stage is seldom included

$$A_{v3} \cong \frac{g_{m8}}{G_L + g_{m8} + g_{ds8} + g_{ds9}} \qquad (3)$$

Where G_L is the load conductance being driven by the buffer stage.
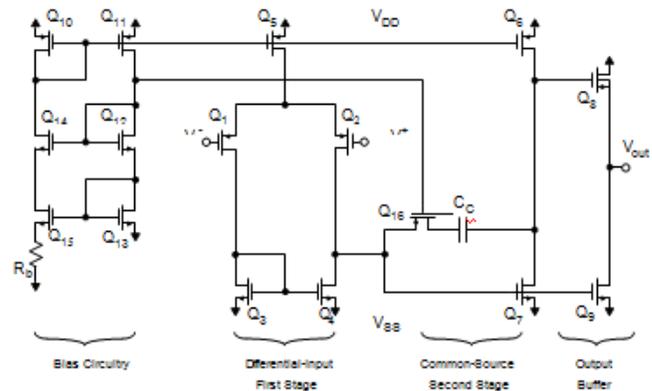
Fig 2 A Two stage Op-amp with second common source stage.

II. COMPENSATION PROCEDURE

Compensation procedure followed is:
a) Start by choosing C_c=5pF arbitrarily.
b) Using SPICE find the frequency where there is a 125° phase shift. Let the gain at this frequency be denoted A'. Also let the frequency be denoted ω_t. This is the frequency that we would like to become the unity-gain frequency of the loop gain[3].
c) Choose a new C_c so that ω_t becomes the unity-gain Frequency of the loop-gain, thus resulting in a 55° phase-margin (and the reason for the choice of 125° used above).This can be achieved by taking C_c according to the equation.
$$C_C = C_c{}' A{}' \qquad (4)$$
It might be necessary to iterate on C_c a couple of times using SPICE.
d) Choose R_C according to

$$R_c = \frac{1}{1.2\,\omega_t} \qquad (5)$$

This choice will increase the unity-gain frequency by about 20%, leaving the zero near to the final resulting unity-gain frequency, which will end up about 15% below the equivalent second pole frequency. The resulting phase margin is approximately –85°. This allows a margin of 5° to account for processing variations without the poles of the closed-loop response becoming real. This choice is also near optimum lead-compensation for almost any Opamp when a resistor is placed in series with the compensation capacitor. It might be necessary to iterate on a couple of times to optimize the phase-margin. However, it should be checked that the gain continues to steadily decrease at frequencies above the new unity-gain frequency, otherwise the transient response can be poor.
This situation sometimes occurs when unexpected zeros at frequencies only slightly greater than are present[1]-[3].

e) If after d), the phase-margin is not adequate, then increase C_C while leaving R_C constant. This will move both ω_t and the lead-zero to lower frequencies, while keeping their ratio approximately constant, thus minimizing the effects of higher frequency poles and zeros which, hopefully, do not also move to lower frequencies.

In most cases, the higher-frequency poles and zeros (except for the lead zero) will not move to significantly-lower frequencies when increasing.

TABLE I
Robust Bias OpAmp Design Procedure [4]

Step 1	$C_C = \dfrac{16kT}{3\omega_u S_n(f)}\left[1 + \dfrac{SR}{\omega_u(V_{HR}^{CM+} + V_{tn})}\right]$
Step 2	$I_{D7} = SR(C_C + C_L)$
Step 3	$L_6 = \sqrt{\dfrac{3\mu_p V_{HR}^{out+} C_C}{2\omega_u C_L \tan\phi_M}}$
Step 4	$W_6 = \dfrac{2SR(C_C + C_L)}{\mu_p C_{OX}\left(V_{HR}^{out+}\right)^2} L_6$
Step 5	$I_{D5} = C_C SR$
Step 6	$(W/L)_{1,2} = \dfrac{\omega_u^2 C_C}{\mu_n C_{OX} SR}$
Step 7	$(W/L)_{5,8} = \dfrac{2SRC_C}{\mu_n C_{OX}\left(V_{HR}^{CM-} - V_{tn} - SR/\omega_u\right)^2}$
Step 8	$(W/L)_7 = \left(\dfrac{C_C + C_L}{C_C}\right)(W/L)_{5,8}$
Step 9	$(W/L)_{3,4} = \dfrac{(W/L)_6}{2(W/L)_7}(W/L)_{5,8}$
Step 10	$I_{D9} = \dfrac{(\tan\phi_M \omega_u C_L)^2\left(C_C + \frac{2}{3}W_9 L_9 C_{OX}\right)^2}{2\mu_n C_{OX}(W_9/L_9)}$
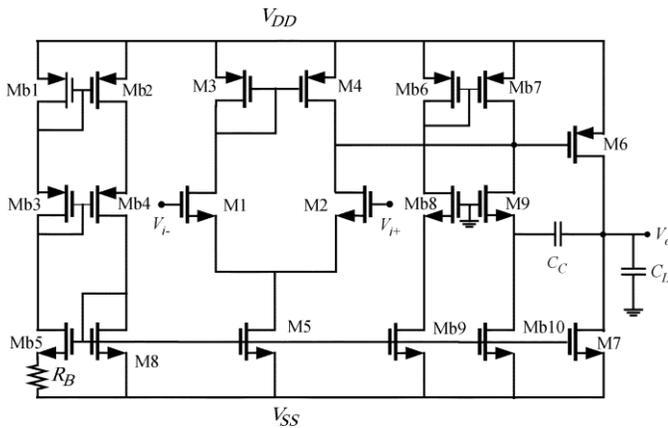
Fig. 3 Opamp with robust bias circuit

TABLE II
Design Parameters For Robust Bias Op Amp

$(W/L)_{1,2}$	.7μm/1μm
$(W/L)_{3,4}$	1.8μm/1μm
$(W/L)_{5,8}$	1.6μm/1μm
$(W/L)_6$	38.6μm/1μm
$(W/L)_7$	17μm/1μm
$(W/L)_9$	27μm/1μm
$(W/L)_{b1-b4}$	1.6μm/1μm
$(W/L)_{b5}$	6.4μm/1μm
$(W/L)_{b6-b7}$	131μ/1μm
$(W/L)_{b8}$	27μm/1μm
$(W/L)_{b9-b10}$	7.4μm/1μm
R_b	32K
C_c	0.5pF

TABLE III
Simulation Results

Opamp with Common Source Stage	Opamp with Current Buffer
GAIN in db	
62	42

III. CONCLUSIONS

Compared to the procedure based upon the nullifying resistor compensation the value of C_C of the proposed procedure [4] can be made much smaller. The wider range of the allowable value of C_C provides a higher flexibility for noise-power tradeoff.

REFERENCES

[1]   P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 2nd ed.   Oxford, U.K.: Oxford, 2002.

[2]   P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design*

Of Analog Integrated Circuits, 4th ed.   New York: Wiley, 2001.

[3]   D. A. Johns and K. Martin, *Analog Integrated Circuit Design*.
New York: Wiley, 1997.

[4]   J.Mahattanakull, "Design Procedure for Two-Stage CMOS Operational Amplifiers Employing Current Buffer," *IEEE Trans. Circuits and Systems-II*, Vol. 52, No. 11, pp. 766-770, Nov.2005. All the simulations are carried out on 0.5 μm Technology using Tanner Tool.

DESIGN OF FIBER REINFORCED PLASTIC LAUNCH TUBE

Srikanth Ananthasagaram[1], V.Gopinath[2]

[1] PG student , [2] Associate Professor
Department of Mechanical Engineering
QIS College Of Engineering &Technology
Ongole,Andhra Pradesh

ABSTRACT

Launch tube is widely used in defense sector to launch the missiles as well as to carry the missiles from one place to another. Use of fiber reinforced plastics reduces the weight of product with out any reduction in the load carrying capacity and stiffness. Because of their the material's high elastic strain energy storage capacity and high strength-to-weight ratio compared with those of steel, fiber reinforced plastics are considered as a materials for construction of launch tubes. Fiber reinforced plastic tubes made of unidirectional carbon fibers embedded in epoxy resin are 65 to70 percent lighter than equivalent steel tubes. Fiber reinforced plastics also exhibit excellent fatigue resistance and durability.

Graphite epoxy composites are widely used in manufacturing launch tubes. Since tubes made of graphite epoxy are often failed at extreme load conditions, there is a continues search for an alternative. In this paper a launch tube is modeled and analyzed for both glass epoxy composites and graphite epoxy composites. The results are tabulated for different orientations of fibers. Both thermal and coupled analyses are carried out. And vonmises stresses are analyzed. Finally it is found that glass epoxy materials possess better properties than graphite epoxy materials and suits well for manufacturing launch tubes.

Key words: Modelling, Analysis, Design, Fiber Reinforced Plastic, Launch Tube

1.0 INTRODUCTION

General purpose of launch tube is to launch the missiles, more over to carry the missiles from one place to another place which are used in defense applications The launch tube contain lugs, the launch lugs are small tubes (Straws), which are attached to the body tube. The launch rail is inserted through these tubes to provide stability to the rocket during launch.

Launch tube is not secured enough inside the launcher, it is pulled out of the launcher and falls back during launching, and the inner part of the launch tube will experience high plume static pressure and plume temperature. Missiles launch from a floating platform. This includes tests of the buoyancy of the launch platform, ease of set-up and use, rate of wind drift, ease of loading missiles into launch tube and erecting on the float gantry, Test of ignition circuits in a wet environment, launch control procedure, and the ability to track and recover missiles. Filament wound launch tubes are just another form of rocket motor cases, with both ends open. Launch tubes employ the same reinforcement materials, design principles, and manufacturing methods.

Launch tubes operate on the same principle as ancient blowgun. A Projectile is inserted into the tube, a gas (air or other) is forced into one end of the tube and the projectile shoots out from the other end.

A General definition of a composite is a synergistic combination of two or more materials, more specifically; the composites referred to here comprise high strength reinforcement in fibrous form, incorporated into and bonded together by matrix, usually a thermosetting polymer. The term fiber reinforced plastics (FRP) is widely used to describe such materials with glass-reinforced plastic (GRP) when the reinforcement is glass fiber. Glass reinforced epoxy (GRE) is used when, as in the case of much composite pipe work, epoxy resin in the matrix.

The use of composite material in the aerospace, rail, marine and civil engineering applications is rapidly increasing. The materials cost economics low weight, high strength and high stiffness, combined with their durability, Means that these materials provide an effective means of achieving design requirements that are driven by consideration weight, Longevity and through life.

1.1 FEA ANALYSIS OF LAUNCH TUBE

Structural analysis is probably the most common application of the finite element method. The term *structural* (or *structure*) implies not only civil engineering structures such as bridges and buildings, but also naval, aeronautical, and mechanical structures such as ship hulls, aircraft bodies, and machine housings, as well as mechanical components such as pistons, machine parts, and tools.

The seven types of structural analyses available in the ANSYS family of products are explained below. The primary unknowns (nodal degrees of freedom) calculated in a structural analysis are *displacements*. Other quantities, such as strains, stresses, and reaction forces, are then derived from the nodal displacements.

Structural analyses are available in the ANSYS Multiphysics, ANSYS Mechanical, ANSYS Structural, and ANSYS Professional programs only.

You can perform the following types of structural analyses. Each of these analysis types is discussed in detail in this manual.

Static Analysis--Used to determine displacements, stresses, etc. under static loading conditions. Both linear and nonlinear static analyses. Nonlinearities can include plasticity, stress stiffening, large deflection, large strain, hyper elasticity, contact surfaces, and creep.

Modal Analysis--Used to calculate the natural frequencies and mode shapes of a structure. Different mode extraction methods are available.

Harmonic Analysis--Used to determine the response of a structure to harmonically time-varying loads.

Transient Dynamic Analysis--Used to determine the response of a structure to arbitrarily time-varying loads. All nonlinearities mentioned under Static Analysis above are allowed.

Spectrum Analysis--An extension of the modal analysis, used to calculate stresses and strains due to a response spectrum or a PSD input (random vibrations).

Buckling Analysis--Used to calculate the buckling loads and determine the buckling mode shape. Both linear (eigenvalue) buckling and nonlinear buckling analyses are possible.

Explicit Dynamic Analysis--This type of structural analysis is only available in the ANSYS LS-DYNA program. ANSYS LS-DYNA provides an interface to the LS-DYNA explicit finite element program. Explicit dynamic analysis is used to calculate fast solutions for large deformation dynamics and complex contact problems

2.0 MODELING OF LAUNCH TUBE

The modeling of the launch tube was done in CATIA software. This model was transferred to ANSYS software through IGES file format.
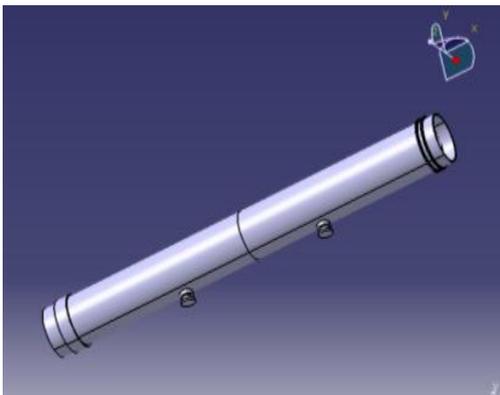
Fig 1: Modeling of the launch tube in CATIA   **Fig 2:** Modeling of the launch tube in ANSYS
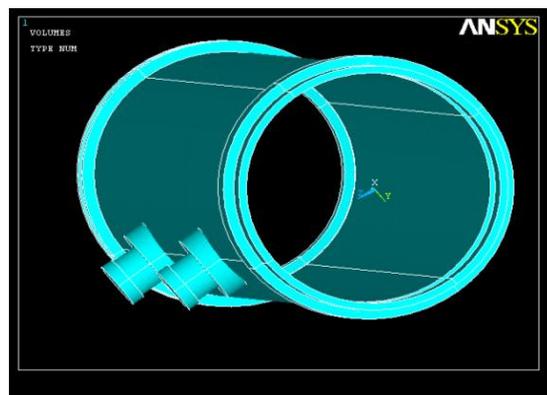
2.1 Material Properties

2.1.1 Physical properties of E Glass fiber

Physical property	Metric	Comment
Density	2.54 - 2.6 g/cc	Independent of length& size of tube

2.1.2 Mechanical properties of E Glass fiber

Tensile Strength, Ultimate	3448 MPa	At 23°C (73°F); Virgin strength. 50-75% variation in finished product; 5310 MPa at -190°C (-310°F); 2620 MPa at 370°C (700°F); 1725 MPa at 540°C (1000°F)
Modulus of Elasticity	72.4 GPa	at 23°C (73°F); 72.3 GPa at 540°C (1000°F)
Poisson's Ratio	0.25	Independent
Shear Modulus	30 GPa	Calculated

2.1.3 Thermal properties of E Glass fiber

Physical property	Metric	English	Comment
CTE, linear 20°C	5 μm/m-°C	2.78 μin/in-°F	
CTE, linear 250°C	5.4 μm/m-°C	3 μin/in-°F	from -30 to 250°C (-20 to480°F)
Heat Capacity	0.81 J/g-°C	0.194 BTU/lb-°F	at 23°C (73°F); 1.03 J/g-°C (0.247 Btu/lbf-°F) at 0°C (390°F)
Thermal Conductivity	1.3 W/m-K	9.02 BTU-in/hr-ft²-°F	
Melting Point	Max 1725 °C	Max 3140 °F	

2.1.4 Mechanical properties of graphite fiber

Modulus of elasticity	207 Gpa
Tensile strength	1035 Mpa
Poisson's Ratio	.25
Shear modulus	2.6 Gpa

2.2 Viewing the results in general post processor.

Failure stresses i.e. vonmises stresses have been seen in general post processor by using contour plot option. In this option again by selecting nodal solution option vonmises stresses have been observed. The stress plot of the launch tube, which is having 6mm thickness and having 45^0 layer orientation angles. Is shown as below.

2.2.1 Graphite

Load applied
Radial pressure = 0.33 MPa
Axial pressure = 2.84625 MPa
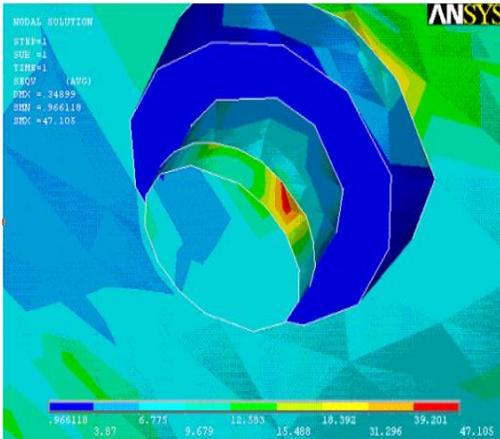
Fig3: Maximum stress plot at mounting lug for 6mm thickness Launch tube having 45^0 orientation angle of graphite fiber
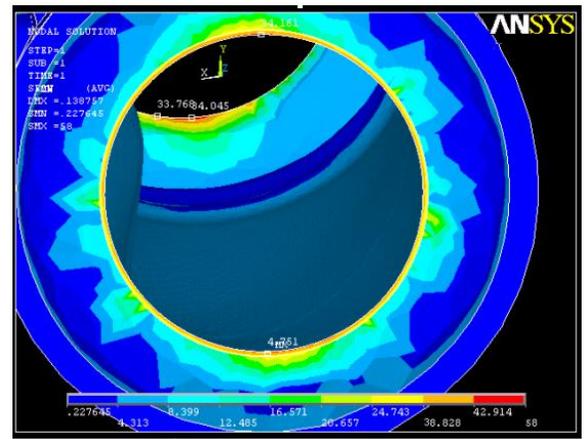
Fig 4: The stresses obtained in theses cases are out of safe limits, so further analysis is done for glass epoxy composite material
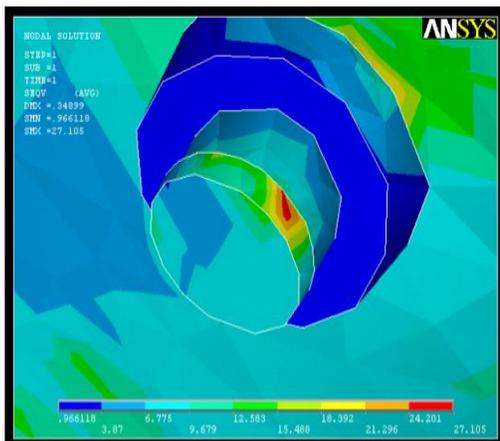
Fig 5: Maximum stress plot at mounting lug for 5mm thickness launch tube having 45^0 layer orientation angle
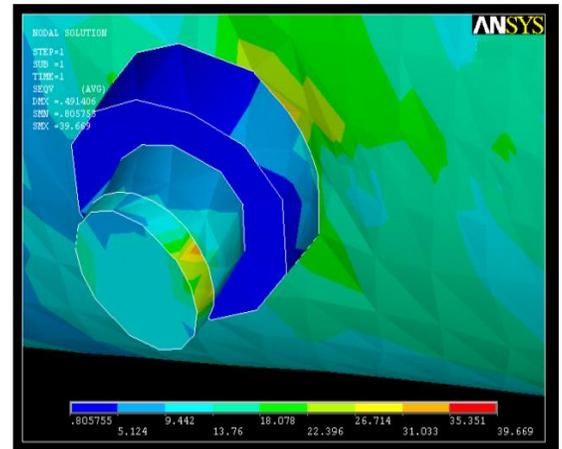
Fig 6 Maximum stress plot at mounting lug for 3mm thickness launch tube having 45^0 layer orientation angle

Table showing different values of vonmises stresses for different layer orientation angles.

	Layer Thickness 6 mm	Layer Thickness 5mm	Layer Thickness 4mm	Layer Thickness 3mm
Layer Orientation Angle 15^0	26.594 Mpa	27.415 Mpa	32.024 Mpa	41.025 Mpa
Layer Orientation Angle 30^0	26.421 Mpa	28.174 Mpa	31.246 Mpa	40.348 Mpa
Layer Orientation Angle 45^0	25.695 Mpa	27.105 Mpa	31.371 Mpa	39.669 Mpa
Layer Orientation Angle 60^0	26.097 Mpa	27.184 Mpa	32.338 Mpa	40.554 Mpa
Layer Orientation Angle 75^0	25.998Mpa	27.286Mpa	32.018Mpa	40.024Mpa

By observing the above values in the table low stress values were obtained for 45^0 layer orientation angle for 6mm, 5mm, and 3mm thickness launch tubes. But where as for 4mm thickness launch tube low stress value was obtained for 30^0 layer orientation angle. So 45^0 layer orientation angle is the preferable angle in manufacturing of the launch tubes.

So maximum of 39.669 MPa have been obtained for 3mm thickness launch tube, which is having 45^0 layer orientation angle. By manufacturing the launch tube with 3mm thickness launch tube and with 45^0 layer orientation angle 38% reduction in both weight and cost have been obtained

3.0 RESULTS AND DISCUSSIONS
3.1 Thermal Analysis

Now thermal analysis has been done in order to find out the thermal stresses. Maximum of 85^0C was applied on the inner surface of the launch tube. Figures regarding thermal analysis are shown below. Maximim of 3.508 MPa has been found out which is with in the allowable limit only.
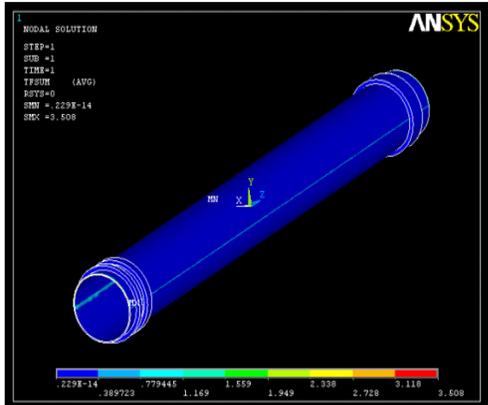
Fig 7: Thermal stress plot of launch tube
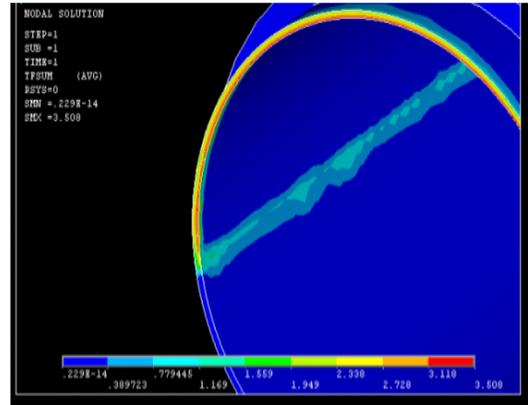
Fig 8: clear representation of thermal stress at side edges of launch tube

3.2 Coupled Field Analysis

Analysis where in the results of one analysis form as input for the other analysis is referred to as coupled field analysis. Due to the complex nature of the physical processes being modeled, it not unusual to conduct coupled analyses as part of a design program. Fluid-structural, fluid-thermal and fluid-acoustic analyses are most common types. Thermal-Structural is the most commonly performed analysis.

Some of the CAE software has the ability to perform the coupled field analysis automatically where as some do not have that capability. However in both cases the Engineer can run one simulation, obtain out put results and apply them as inputs for the other analysis.

The results of thermal analysis were given as input in the coupled field analysis. That is nothing but applying both pressure and temperatures simultaneously. Failure stresses in X, Y and Z directions were found out as 18.516MPa, 2.965MPa and 6.041MPa respectively. Figures of failure stress plots in coupled field analysis are
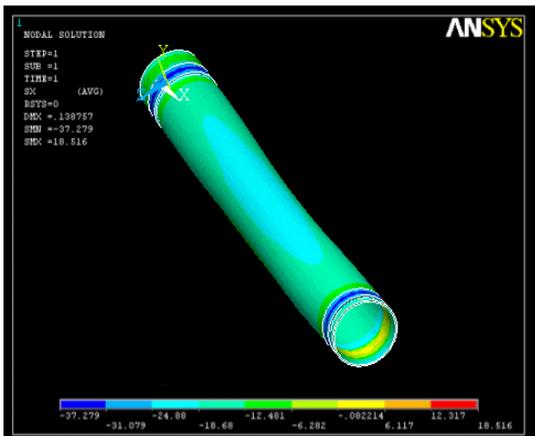
Fig 9: stress plot in coupled field analysis in X direction
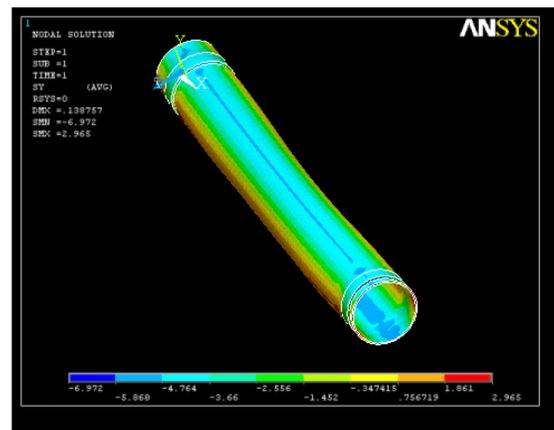
Fig 10: stress plot in y direction in coupled field analysis
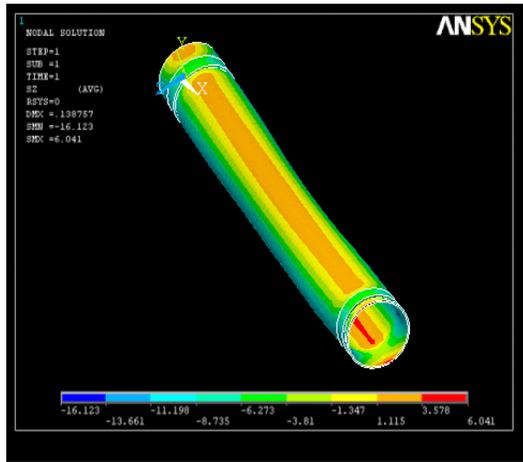
Fig 9.4 stress plot in coupled field analysis in Z direction
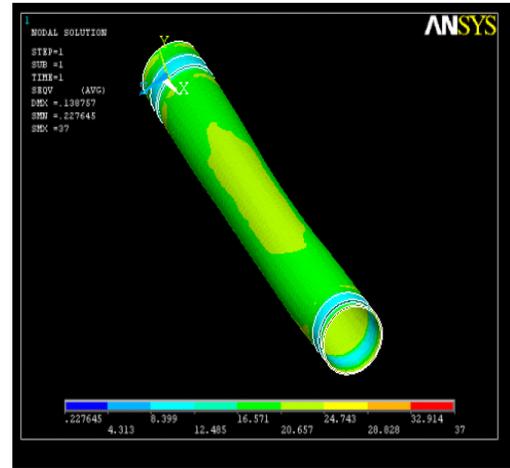
Fig 9.5 vonmises stress plot in coupled field analysis

So obtained maximum vonmises stress is 37 MPa in the coupled field analysis. The design stress limit is 72.426 MPa for the given pressure and temperature .so the design of launch tube which is having 3 mm thickness is the safest design, because the obtained stress value is with in the allowable limit only.

4.0 CONCLUSIONS
(i) Structural Analysis

By observing the values from the results table, low stress values were obtained for 45^0 layer orientation angle for 6mm, 5mm, and 3mm thickness launch tubes. But where as for 4mm thickness launch tube low stress value was obtained for 30^0 layer orientation angle. So 45^0 layer orientation angle is the preferable angle in manufacturing of the launch tubes.

So maximum of 39.669 MPa have been obtained for 3mm thickness launch tube, which is having 45^0 layer orientation angle. But the allowable design stress limit is 72.426MPa.s the obtained stress value is with in the allowable limit only. So manufacturing of the launch tube can be done with minimum amount of material, which leads to low weight and low cost. And more over 38% reductions in weight and cost have been obtained.

(ii) Thermal Analysis

For FRP materials temperature limits for most common applications are 320^0k and for some applications it is up to 340^0k. And for some special cases it is up to 400^0k.in this project applied temperature is 358^0k.The design stress limit for the given temperature is 72.426 MPa, the maximum stress obtained in this thermal analysis is 3.509 Mpa, which is with in the allowable limit only.

(iii) Coupled Field Analysis

For combined temperature and pressure loads the maximum vonmises stress obtained in coupled field analysis = 37 MPa .and the design stress limit is 72.426 MPa for the given pressure and temperature the obtained stress value is with in the allowable stress limit only.

In the above three analyses the obtained stresses are with in the allowable stress limits only. Developed design of launch tube which is having 3 mm thickness in this project can be with stand for the application of combined pressure and temperature loads and more over it can be manufacturability also.

5.0 REFERENCES

[1]   Peters, S.T*; "Handbook of Composites"*; *Chapman and Hall London*; 2nd Edition*;* 1998.

[2]   Geoff Eckold*;* "Design and Manufacture of Composite Structures"*; Jaico publishing House;* 1995.

[3]   Agarwal, B. D and Brootman, L. J*;* Analysis and performance of Fiber Reinforced Composites*; Wiley publishers;* 2nd Edition*;* 1990. (Xerox copy)

[4]   Jones, R. M*;* Mechanics of Composite Materials*; Mc Graw Hill Book Co.;* 1975.

[5]   Mallick, P. K*;* Composite Engineering Handbook*;* 1997

[6]   Vasiliev, V. V and Morozov, E. V*;* Mechanics and Analysis of Composite Materials*; Elsevier Science;* 2001.

[7]   Kumar, S. and Pradhan, B*;* Finite Element Analysis of Low-Velocity Impact Damage in Composite Laminates*; Journal of Reinforced Plastics and Composites*; Vol.19*;* No. 04/2000.

[8]   Heintz, Chris*;* Article on Wood, Aluminum, Steel and Composites; Series on Light Aircraft Design and Construction*; Published in EAA Light Plane World Magazine;* Dec. 1985.

[9]   Faruk Omar*;* Article on Natural and Wood Fiber Reinforced Polymer Composites*; Institute or Material Science;* University of Kassel*;* Germany.

[10]   Clare Dominic*;* A Survey of Materials and their Properties in terms of Design and Technology.

[11]   J. Steven Mayes; Andrew C. Hansen: Multicontinuum Failure Analysis of Composite Structural Laminates

[12]   Filament Winding Material Properties*;* C-K Composites*,* Mount Pleasant, PA.

[13]   Dr. s. Li*;* Technical paper on, Failure criteria of composites; November 2002

[14]   Federico Paris; A study of failure criteria of fibrous composite material; March 2001.

[15]   M. Sonnen, C. Laval, A. Seifert; Technical paper on; new ways in analytical calculation of laminates and composite structures; October 2003

[16] zein kiewicz /O.C,"The Finite Element Method", Third Edition, MeGraw-hill Book Company, 1997

Seismic Performance Enhancement Methodology for Framed Structures using Supplemental Damping

K. Sathish Kumar[1], R. Sundararajan[2], C. Antony Jeyasehar[3], T. Radha Krishnan[4] & K. Muthumani[5]

[1] (Senior Principal Scientist, ASTaR Laboratory, CSIR-Structural Engineering Research Centre, Taramani, Chennai, India)
[2] (Professor, Department of Structural Engineering, Government College of Technology, Coimbatore, India)
[3] (Professor, Department of Civil and Structural Engineering, Annamalai University, Annamalai Nagar, India)
[4] (M.E., Student, Department of Structural Engineering, Government College of Technology, Coimbatore, India)
[5] (Chief Scientist, ASTaR Laboratory, CSIR-Structural Engineering Research Centre, Taramani, Chennai, India)

Abstract
Supplemental damping through passive energy dissipation (PED) devices is often used for enhancing the seismic performance of a seismically deficient structure to reduce the seismic response under earthquake loading. Such PED devices are normally incorporated within the frame structure between adjacent floors through different bracing schemes like diagonal and chevron, so that they efficiently enhance the overall energy dissipation ability of the seismically deficient frame structure in the loading direction. These PED devices function based on the large and stable energy dissipation obtained using energy dissipation mechanisms like visco-elastic and elasto-plastic. This paper presents a methodology based on the direct displacement based design (DBD) for designing PED devices for providing supplemental damping to enhance the energy dissipation ability of multi-storey frames subjected to earthquake loading.

Keywords - Seismic performance enhancement, seismic retrofitting, displacement based design, supplemental damping, passive energy dissipation device

1. INTRODUCTION

Recent damaging earthquakes provided powerful reminders of how vulnerable we all are to the forces of nature. Even in an advanced industrial nation, our built environment is still quite susceptible to natural disasters. Consequently, one of the principal current challenges in structural engineering concerns the development of innovative design concepts to better protect structures, along with their occupants and contents, from the damaging effects of destructive environmental forces due to earthquakes. The traditional approach to seismic design has been based on providing a combination of strength and ductility to resist the imposed loads. For major earthquakes, the structural design engineer relies upon the inherent ductility of structure to prevent catastrophic failure, while accepting a certain level of damage. In this traditional seismic design, acceptable performance of a structure during an earthquake is based on the lateral force resisting framed system being able to absorb and dissipate energy in a stable manner for a large number of cycles. Energy dissipation occurs in specially detailed ductile plastic hinge regions of beams and column bases which also form part of the gravity load carrying system. Plastic hinges are regions of concentrated damage to the gravity frame which often is irreparable. Nevertheless, this design approach is acceptable because of economic considerations provided, of course, that structural collapse is prevented and life safety is ensured. Sometimes, situations exist in which this traditional seismic design approach is not applicable. When a structure must remain functional after an earthquake, as the case of lifeline structures, the conventional seismic design approach is inappropriate. For such cases, the structure may be designed with sufficient strength so that inelastic action is either prevented or is minimal; an approach that is very costly. Moreover, in such structures, special precautions need to be taken in safeguarding against damage or failure of important secondary systems which are needed for continuing serviceability. But this draw back can be mitigated, and perhaps eliminated, if the earthquake-induced energy is dissipated in supplemental damping devices placed in parallel with the gravity load resisting system. The new approach for improving seismic performance and damage control is that of passive energy dissipation (PED) systems. This strategy is attractive for two primary reasons:

1. Damage due to the gravity load resisting system is substantially reduced, leading to major reduction in post earthquake repair costs.
2. Earthquake damaged PED devices can be easily replaced without the need to shore the gravity framing.

Alternate seismic performance enhancement strategies [1] have been developed which incorporate earthquake protective systems in the structure. In these systems, mechanical devices are incorporated into the frame of the structure to dissipate energy throughout the height of the structure. The means by which energy is dissipated is either yielding of mild steel,

sliding friction, motion of a piston or a plate within a viscous fluid, orifice action of fluid or visco-elastic action in polymeric materials. In addition to increasing the energy dissipation capacity per unit drift of a structure, some energy dissipation systems also increase the strength and stiffness. Such systems include the following types of energy dissipation mechanisms: yielding, extrusion, friction, viscous and visco-elastic action.

2. MECHANISM OF SUPPLEMENTAL DAMPING

Fig. 1(a) and Fig. 1(b) show the pushover curves of a linearly elastic frame and yielding frame which is essentially a plot of base shear vs. top floor displacement.   Similarly, the corresponding force displacement hysteretic loops depict linear behavior and limited ability to absorb energy.

Consider the case when energy-dissipating devices are added to the frame, it is assumed that the connection details of the devices are such that neither inelastic action nor damage occurs in the frame at the points of attachment during seismic excitation. It is also assumed that the design of the energy dissipation system is such that it functions properly and dissipates energy throughout the height of the frame. The ability of the frame to dissipate energy is substantially increased as demonstrated in the force-displacement hysteretic loops of the frame. Accordingly, the frame undergoes considerably reduced amplitude of vibration in comparison to the frame without the energy dissipation system under the same earthquake motion. While the energy dissipation system can achieve a considerable reduction in the displacement response, it can also achieve a reduction in the total force exerted on the structure. In general, reduction in force will not be as much as reduction in displacement which is due to the increased strength or increased stiffness provided by the energy dissipation system. Comparable reductions in displacement and force can be achieved with systems that do not increase the strength or stiffness of the structure to which they are attached.

3. MODELING OF PED DEVICES

For analysis of structures with PED devices, various mathematical modeling techniques have been developed. Various models with increased complexity are reviewed in Reinhorn et al., (1995) [2] for PED devices of viscous type. Constantinou and Symans (1993) [3] showed that the Maxwell model is adequate to capture the frequency dependence of the viscous PED device. It is also shown that, below a cut off frequency of approximately 4 Hz, the model can be further simplified into a purely viscous dashpot model. It is stated in FEMA-273 [4] and FEMA-274 [5] that the damping force of a viscous PED device can be modeled to be proportional to the velocity with a constant exponent ranging between 0.2 and 2.0. In preliminary analysis and design stages, the velocity exponent of 1.0 is recommended for simplicity. In this study, based on those references, the behavior of viscous PED device is modeled by a linear dashpot.
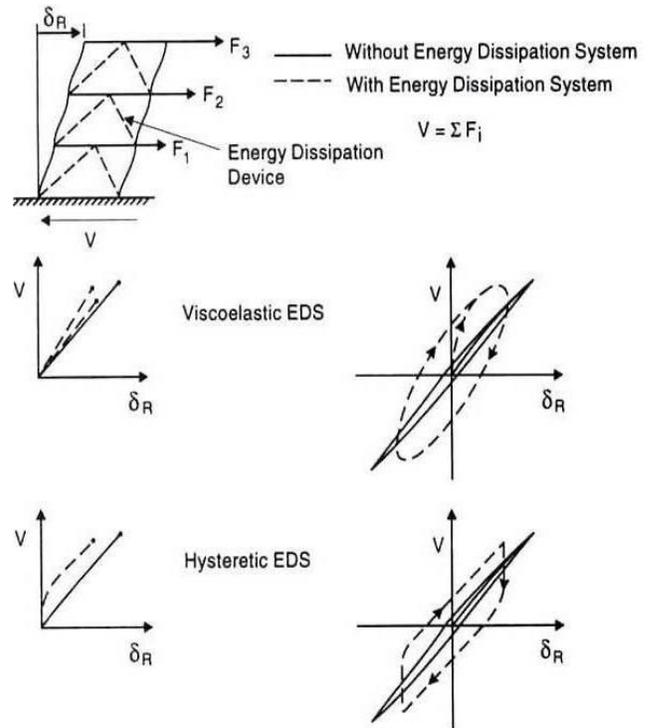
Fig.1 (a) Pushover curves and force-displacement hysteretic curves of an elastic structure without and with passive energy dissipation devices
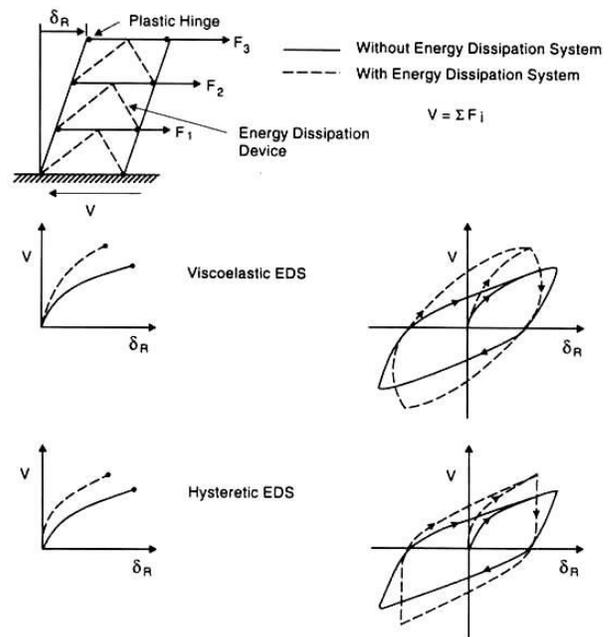
Fig.1 (b) Pushover curves and force-displacement hysteretic curves of a yielding structure having proper plastic hinge formation, without and with passive energy dissipation devices

A typical visco-elastic PED device consists of thin layers of visco-elastic material bonded between steel plates. In practice, the dynamic behavior of visco-elastic PED device is generally represented by a spring and a dashpot connected in parallel [6]. For the linear spring-dashpot representation of the visco-elastic PED device, the stiffness K_d and the damping coefficient C_d are obtained as follows:

$$K_d = \frac{G'(\omega)A}{t}$$
$$C_d = \frac{G''(\omega)A}{\omega t} \qquad (1)$$

Where, $G'(\omega)$ and $G''(\omega)$ are the storage shear modulus and loss shear modulus respectively; A and t are total shear area and the thickness of the material respectively; and ω is forcing frequency for which the fundamental natural frequency of the structure is generally utilized in time domain analysis. With this spring-damper idealization, the dynamic system matrices of the structure with added visco-elastic PED devices can be constructed by superposing the damper properties to the stiffness and damping matrices of the structure. Fig. 2 represents the mathematical models of viscous and visco-elastic PED devices employed in this study.
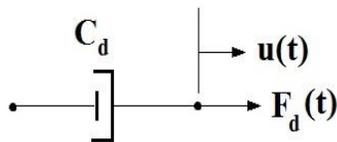
Fig. 2(a) Mathematical model representing viscous PED device
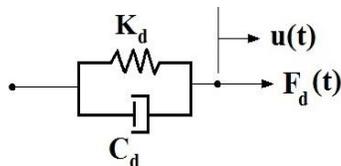
Fig. 2(b) Mathematical model representing visco-elastic PED device

4. PERFORMANCE EVALUATION USING DISPLACEMENT SPECTRUM AND CAPACITY CURVE

The direct displacement based design (DBD), which focuses on displacement as the key design parameter, is considered to be an effective method for implementing performance based seismic design utilizing deformation capacity and ductile detailing standards. In the present study, the general procedure of the DBD documented in the SEAOC Blue Book [7] is applied in reverse order for evaluation of seismic performance of an existing structure. In principle, the proposed analysis procedures are similar to the capacity spectrum method [5],[8],[9] in that performance point is determined as a location where the displacements demand of the earthquake becomes equal to the plastic deformation capacity of the structure. The difference is on the use of displacement spectrum instead of the so called acceleration displacement response spectrum (ADRS). Therefore, the extra work required for transforming the capacity and demand curves to ADRS format can be avoided. Although this may not be a significant improvement, it has the advantage of maintaining consistence with the proposed design procedure for supplemental dampers. Two nonlinear static analysis procedures, the step by step and the graphical procedure, which correspond to the nonlinear static procedures A and B of ATC-40 [8] respectively, are proposed for seismic performance evaluation of structures (without PED devices). The two procedures are summarized as in the following sub-sections:
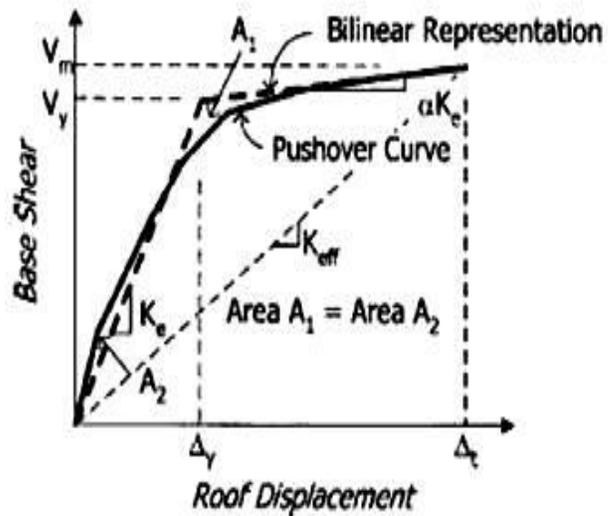
Fig. 3 Bilinear representation of a capacity (pushover) curve

4.1 Step by step procedure

1. Obtain base shear versus roof storey displacement capacity curve for the frame structure from pushover analysis.
2. Approximate the capacity curve by bilinear lines based on equal energy concept (area A_1 = area A_2), and determine the quantities such as effective elastic stiffness K_e, elastic natural period T_e, base shear at yield V_y, yield displacement Δ_y and post-yield stiffness ratio α (Fig.3).

3. Transform the roof storey displacement coordinate into pseudo-displacement coordinate S_d using the following relation:

$$S_d = \frac{\Delta_R}{\Gamma \phi_R} \qquad (2)$$

Where, Δ_R is the roof displacement and Γ and ϕ_R is the modal participation factor and the roof storey component of the fundamental mode respectively. This process corresponds to the transformation of the structure into an equivalent single degree of freedom (SDOF) structure.

4. Assume the first trial value for the maximum displacement S_{dm} of the equivalent structure, and determine the ductility factor $\mu = S_{dm}/S_{dy}$. The equivalent damping ratio ξ_{eq} can be obtained as:

$$\xi_{eq} = \frac{2(\mu-1)(1-\alpha)}{\pi\mu(1+\alpha\mu-\alpha)} \qquad (3)$$

Then, the effective damping for the structure can be obtained as the sum of the equivalent damping and the inherent damping of the structure:

$$\xi_{eff} = \xi_{eq} + \xi_i \qquad (4)$$

Where, ξ_i is the inherent damping for which 5% of critical damping is generally utilized. Also, the effective period T_{eff} corresponding to the maximum displacement can be obtained as:

$$T_{eff} = T_e \sqrt{\frac{\mu}{1+\alpha\mu-\alpha}} \qquad (5)$$

Where, T_e is the fundamental period of the structure.

5. Construct the displacement response spectrum for design earthquake using the effective damping obtained in the previous step, and read from the spectrum the next trial value for the maximum displacement S_{dm} corresponding to the effective period T_{eff}.

6. Repeat the process from step 4 using the maximum displacement computed in the above step. Once the maximum displacement S_{dm} converges, then convert it into the maximum roof displacement using equation 2.

7. Carry out pushover analysis until the roof displacement reaches the maximum value computed above to estimate the maximum inter-storey drifts.

4.2 Graphical procedure

1. Steps 1 & 2: The same as those of the step by step procedure.
2. Step 3: Draw displacement response spectra with various damping ratios.
3. Step 4: For a series of ductility ratios, obtain maximum displacements ($S_{dm} = \mu . S_{dy}$), effective periods $T_{eff} (\mu)$ [Eq.5] and effective damping ratios (ξ_{eff}) [equations.3 and 4].
4. Step 5: Find out the point at which the effective damping ratio corresponding to a ductility ratio, obtain in step 4, is equal to the equivalent damping ratio of a displacement spectrum crossing the point [$T_{eff}(\mu)$, $S_{dm}(\mu)$].
5. Step 6: Convert the maximum displacement computed in the above step into the maximum roof displacement, and carry out pushover analysis until the roof displacement reaches the maximum value computed above to estimate the maximum inter-storey drifts.

5. DESIGN PROCEDURE FOR PED DEVICES

If the maximum storey drift of a structure subjected to a code-specified earthquake load exceeds the desired performance level, the structure needs to be retrofitted. Among the various methods for seismic retrofit, this study focuses on increasing damping to decrease earthquake induced structural responses. To this end, a procedure for estimating the amount of supplemental damping required to satisfy the given performance objective is proposed. The basic idea is to compute the required damping from the difference between the total effective damping needed to meet the target displacement and the equivalent damping provided by the structure at the target displacement.

5.1 Required damping to meet target displacement

The damping ratio of the displacement response spectrum that intersects the point of the target displacement S_{dt} on the displacement ordinate (vertical axis) and the effective period T_{eff} on the period ordinate (horizontal axis) corresponds to the total effective damping ξ_{eff} for the structure to retain to meet the performance objective. For structure with supplemental dampers, the total effective damping is composed of the three components: inherent viscous damping ξ_i, equivalent damping of the structure contributed from inelastic deformation of the structural members ξ_{eq} and the damping required to be added by the PED devices ξ_d. The equivalent damping of the structure is obtained from the following equations [4]:

$$\xi_{eq} = \frac{1}{4\pi}\frac{E_{DS}}{E_S} = \frac{V_y S_{dt} - S_{dy} V_t}{\pi V_t S_{dt}} \qquad (6a)$$

for Viscous PED device

$$\xi_{eq} = \frac{V_{yd} S_{dt} - S_{dy} V_{td}}{\pi V_{td} S_{dt}} \quad (6b)$$

for Visco-elastic PED devices

Where, $V_{yd} = V_y + K_d S_{dy}$, $V_{td} = V_t + K_d S_{dt}$ and E_s and E_{DS} are the stored potential energy in the structure and the energy dissipated by hysteretic behavior of the structural members in the retrofitted structure respectively. Tsopelas et al., (1997) [10] provides the contribution of the added damping to the total effective damping as $(\xi_d \cdot T_{eff})/T_e$, where ξ_d is the supplemental damping ratio. Then the required supplemental damping can be computed from the following equation:

$$\xi_d = (\xi_{eff} - \xi_{eq} - \xi_i)\frac{T_e}{T_{eff}} \quad (7)$$

Where, the total effective damping and the equivalent damping can be obtained from the displacement response spectrum and from equation 6 respectively.

5.2 Storey-wise distribution of PED devices
In multi-storey frame structures, the supplemental damping computed in the equivalent SDOF system using equation 7 should be distributed throughout the stories of the original structure in such a way that the damping ratio for the fundamental mode becomes the required supplemental damping ξ_d. For this purpose, the expression for equivalent damping (equation 6) is used again except that the energy dissipated by the PED device E_{DV} is used in the numerator instead of E_{DS}

$$\xi_{eq} = \frac{1}{4\pi}\frac{E_{DV}}{E_S} \quad (8)$$

If the PED devices are placed as diagonal members with the inclination θ, then, the energy dissipated by the PED devices can be expressed as follows [4]:

$$E_{DV} = \frac{2\pi^2}{T_{eff.d}}\sum_{i=1}^{N} C_{di}\cos^2\theta_i(\Delta_i - \Delta_{i-1})^2 \quad (9)$$

Where, $T_{eff.d}$ is the secant period of the retrofitted structure; C_{di} and Δ_i are the damping coefficient and the maximum lateral displacement of the ith storey respectively, and N is the number

of storey. The potential energy stored in the multi-storey structure can be expressed as follows:

$$E_S = \frac{2\pi^2}{T_{eff.d}}\sum_{i=1}^{N} m_i \Delta_i^2 \quad (10)$$

$$T_{eff.d} = 2\pi\sqrt{\frac{M * S_{dt}}{V_y(1+\alpha\mu-\alpha)}} \quad (11a)$$

for Viscous PED devices

$$T_{eff.d} = 2\pi\sqrt{\frac{M * S_{dt}}{V_y(1+\alpha\mu-\alpha) + K_d S_{dt}}} \quad (11b)$$

for Visco-elastic PED devices

Where, $M*$ is the effective modal mass and m_i is the mass of the ith storey. By substituting equation 9 and 10 into equation 8, the damping ratio contributed from the PED devices can be expressed as:

$$\xi_d = \frac{1}{4\pi}\frac{T_{eff.d}\sum_{i=1}^{N} C_{di}\cos^2\theta_i(\Delta_i - \Delta_{i-1})^2}{\sum_{i-1}^{N} m_i \Delta_i^2} \quad (12)$$

In equation 12, the left hand side of the equation ξ_d is obtained from equation 7 in the equivalent SDOF system. For viscous device, the damping coefficient of the damper device in the ith storey C_{di} can be determined in equation 12, whereas for visco-elastic device, both C_{di} and K_{di} are the variables that should be determined. This can be done by using the relation $K_d = (G'/G'')\omega C_d$ obtained from equation 1. The simplest case is to assume that the PED devices in all storeys have the same capacity, and the damping coefficient in this case can be obtained from equation 12 as:

$$C_d = \frac{1}{4\pi}\frac{4\pi\xi_d\sum_{i=1}^{N} m_i \Delta_i^2}{T_{eff.d}\sum_{i=1}^{N}\cos^2\theta_i(\Delta_i - \Delta_{i-1})^2} \quad (13)$$

In this stage, however, the maximum storey displacements, except for the top-storey displacement given as performance limit state, are known. Therefore, the configuration for lateral storey drifts Δ_i needs to be assumed in equations 12 and 13. A simple case is to assume that the maximum storey drifts are proportional to the fundamental mode shape or to the pushover curve. The storey-wise distribution pattern for the PED devices also needs to be assumed. For viscous dampers, the

design process ends here. However, for PED devices with stiffness such as visco-elastic or hysteretic dampers, iteration is required, because the added PED devices increase system stiffness. In that case, the capacity curve of the system needs to be redrawn considering added PED devices, and the process is repeated until convergence.

6. DESIGN PROCEDURE FOR PED DEVICE SCHEME

The proposed procedure to design supplemental dampers for performance based seismic retrofit of existing structures can be summarized in the following steps:

1. Carry out eigen value analysis of the structure to obtain natural periods and mode shapes. Using the mode shapes, perform pushover analysis to obtain top storey versus base shear curve, and transform the pushover curve into a capacity curve using equation 2. Idealize the curve into a bilinear shape, and read the yield displacement S_{dy}.

2. Decide a desired target roof displacement, and transform it into the target value in the equivalent SDOF system S_{dt}. Obtain ductility ratio $\mu S_{dt}/S_{dy}$, the effective period T_{eff} (equation 5) and the equivalent damping ξ_{eq} (equation 6) at the target displacement.

3. Find out the effective damping ratio corresponding to the displacement response spectrum that crosses the point of the target displacement and the effective period. This corresponds to the total demand on damping imposed by the earthquake. It would be more convenient to start the procedure with response spectra with various damping ratios.

4. Compute the required damping for supplemental dampers from equation 7.

5. The required damping is distributed throughout the storeys using Equation 12. The size of PED device in each storey is designed based on the required damping allocated to the storey.

6. For structures retrofitted with visco-elastic PED devices, carry out eigen value analysis and redraw the capacity curve of the structure using the newly obtained mode shape, and repeat step 1 until convergence.

7. Check whether the structural members, especially columns, can resist the additional axial and shear forces imposed by PED devices.

7. SUMMARY AND CONCLUSIONS

The general procedure of the direct displacement based design (DBD) documented in the SEAOC Blue Book is applied in reverse order for evaluating the seismic performance of an existing structure. Based on which a methodology is presented for designing PED devices of viscous and visco-elastic types for providing supplemental damping to enhance the energy dissipation ability of multi-storey frames subjected to earthquake loading.

REFERENCES

[1] J.M. Kelly, R.I. Skinner, and A.J Heine, Mechanism of Energy Absorption in Special Devices for Use in Earthquake Resistant Structures, *Bull, N.Z., Soc. Earthquake Engineering, 5(3)*, 1972, 63-88.

[2] A.M. Reinhorn, C. Lie, and M.C. Constantinou, *Experimental and Analytical Investigation of Seismic Retrofit of Structures with Supplemental Damping, Part I: Fluid Viscous Damping Devices* (Technical Report, NCEER-95-0001.National Center for Earthquake Engineering Research, Buffalo, New York, 1995).

[3] M.C Constantinou, and M.D. Syman, Experimental Study of Seismic Response of Buildings with Supplemental Fluid Dampers, *Struct. Des. Tall Buildings, 2(2)*, 1993, 93-132.

[4] Federal Emergency Management Agency (FEMA), *NEHRP Guidelines for the Seismic Rehabilitation of Buildings, FEMA-27* (prepared by the Applied Technology Council for the Building Seismic Safety Council, Washington, D.C.,1997a).

[5] Federal Emergency Management Agency (FEMA), *NEHRP Commentary on the Guidelines for the Seismic Rehabilitation of Buildings, FEMA-274* (prepared by the Applied Technology Council for the Building Seismic Safety Council, Washington, D.C. 1997b).

[6] T.T. Soong, and G.F. Dargus, *Passive Energy Dissipation Systems in Structural Engineering* (Wiley, New York, 1997).

[7] Structural Engineers Association of California (SEAOC), *Recommended Lateral Force Requirements and Commentary* (Appendix I, Sacramento, California, 1999).

[8] Applied Technology Council (ATC), *Seismic Evaluation and Retrofit of Concrete Buildings,* (ATC-40, Redwood City, California, 1996).

[9] S.A. Freeman, Development and Use of Capacity Spectrum Method, *Proceedings of 6th National Conference on Earthquake Engineering*, Seatle, 1998.

[10] P. Tsopelas, M.C. Constantinou, C.A. Kircher, and A.S. Whittaker, *Evaluation of Simplified Method of Analysis for Yielding Structures* (Technical Report: NCEER-97-0012, National Center for Earthquake Engineering Research, State University of New York at Buffalo, Buffalo, New York. 1997).

Stability of System matrix via Gerschgorin circles

T.D.Roopamala[1] S.K.Katti[2]
*Department of computer science and eng / S.J.CE, INDIA
** Research supervisor, JSS Research foundation, S.J.C.E INDIA

ABSTRACT
In this paper the stability of the system can be analyzed graphically using Gerschgorin circle theorem. Analytically it has been proved that if the left Gerschgorin bound are very much greater than the right Gerschgorin bound and the trace of the matrix is equal to length of left Gerschgorin bound then there is no eigenvalues on the RHS of s-plane.

Keywords - Eigenvalues, Gerschgorin *bound Gerschgorin circles, stability, and trace*

I. INTRODUCTION

The concept of stability plays very important role in the analysis of the system. In literature there is various methods to find the stability of the system. Given a characteristic polynomial when all the co-efficient of the characteristic polynomial are positive by Routh stability criterion [1] we have to construct the Routh table and in the first column of the Routh array, if there exist change in sign then the system is said to be unstable. This requires the computation of the characteristic polynomial from the system matrix which takes lots of computation and constructing the Routh table requires computation. In this paper attempts have been made to find whether the system is stable graphically. Given a system matrix of order (nxn), than we draw the Gerschgorin circles [2] of the matrix. If the length of the Gerschgorin bound is more on the left hand side of the s-plane and if the trace of the matrix is equal to length left bound, than using Gerschgorin theorem [2] it has proved that the system is stable . This approach does not require any computation.

II. MATHEMATICAL ANALYSIS:

Given   a = left bound.

                T = trace.

                b = right bound.

Also a >> b   a = T is known

To prove that: There exists no eigenvalues on the right half of the of s-plane which implies that the system is stable.

Determination of the bounds of the using Gerschgorin's theorem:

Consider a system matrix $[A]_{nxn} \in R^{nxn}$

For Row wise circles:-

$$L_{rk} = a_{kk} - \sum_{\substack{k=1 \\ i \neq k}}^{n} a_{ik} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (1)$$

$$L_r = \min(L_{rk}) \quad k = 1, 2, 3, ..., n \qquad \rightarrow (2)$$

$$R_{rk} = a_{kk} - \sum_{\substack{k=1 \\ i \neq k}}^{n} a_{ik} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (3)$$

$$R_c = \max(R_{ck}) \quad k = 1, 2, 3, ..., n \qquad \rightarrow (4)$$

L_{rk} = is the left bound for each row-wise circle

L_r = extreme left bound for row-wise circle.

R_{rk} = is the right bound for each row-wise circle.

R_c = extreme right bound for row-wise circle.

For Column wise circles:-

$$L_{ck} = a_{kk} - \sum_{\substack{k=1 \\ j \neq k}}^{n} a_{kj} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (5)$$

$$L_c = \min R_{ck} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (6)$$

$$R_{ck} = a_{kk} - \sum_{\substack{k=1 \\ j \neq k}}^{n} a_{kj} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (7)$$

$$R_r = \max R_{rk} \quad k = 1, 2, 3, ..., n \qquad \rightarrow (8)$$

L_{ck} = is the left bound for each column wise circle.

L_c = extreme left bound for column-wise circle.

R_{ck} = is the right bound for each column-wise circle.

R_c = extreme right bound for column -wise circle.

To obtain the left bound:

If $L_c , L_r < 0$ then the left bound

$a = \max(L_c, L_r)$

f $L_c , L_r > 0$ then the left bound

$a = \min(L_c, L_r)$

If $L_c \leq 0, L_r > 0$ then the left bound $a = L_c$

If $L_r \leq 0 , L_c > 0$ then the left bound

$a = L_r$

To obtain the right bound:

If R_c , $R_r < 0$ then the left bound $b = \max(R_c, R_r)$

If $R_c, R_r > 0$ then the left bound

$b = \min(R_c, R_r)$

If $R_c \leq 0$, $R_c > 0$ then the left bound

$b = R_r$

If $R_r \leq 0$, $R_c > 0$ then the left bound

$b = R_c$

Since a >> b , let L_c , $L_r < 0$ then the left bound

$a = \max(L_c, L_r)$

Let us suppose that $a = L_c$ -Extreme left bound of the column wise circles.

If $R_c, R_r > 0$ then the left bound

$b = \min(R_c, R_r)$

Let us suppose that $b = R_c$ - Extreme right bound of the column wise circles.

Given a >> b, and a = T

$$\text{trace} = \sum_{i=1}^{k} a_{ii} - sum\, of\; principle\, diagonal\, elements$$

$$\text{trace} = \sum_{i=1}^{k} \lambda_{ii} - sum\, of\; eigenvalues$$

Since $a = L_c$ $b = R_c$ Also given that a >> b

$L_c \gg R_c$ and $L_c = trace < 0$

$$\min\left(a_{kk} - \sum_{\substack{k=1 \\ i \neq k}}^{n} a_{ik}\right) < 0, k = 1, 2, 3, ..., n$$

➔ All centers of the circles are on the left half of the s-plane. (Since a >> b)

➔ There exist no circles with the centre on the RHS s-plane. Hence there exists no eigenvalues on the RHS of s-plane.

III. EXAMPLE

$$\begin{pmatrix} -3 & 1 & 1 \\ 3 & -1 & -1 \\ 1 & 1 & -1 \end{pmatrix} \qquad \rightarrow (9)$$

Gerschgorin circles of the above matrix is
Eigenvalues of the above matrix are

$\lambda = 0$
$\lambda = -1$
$\lambda = -4$

The max left is −5
The max right is 1
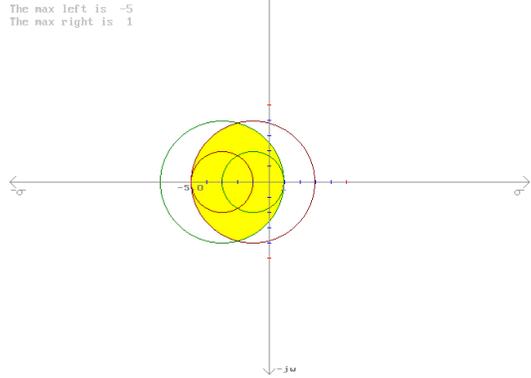
Gerschgorin bound [- 6, 1]

Conclusion:

Hence given a system matrix A of any order , with the condition that the left bound is very much greater than the right bound and also the trace of the matrix is equal to the length of the left bound , than the system does not contain any eigenvalues on the right hand side of s-plane. This is applicable only for the few class of matrices. The advantage of this graphical approach is it requires no computation. Also there exist no eigenvalue on the right half of s-plane. By observing the Gerschgorin circles the stability can be identified.

REFERENCES

[1] I.J.Nagrath, M.Gopal, "Control systems Engineering", Published by New Age International (p) Ltd., Fifth Edition-2007. Reprint-2008.

[2] M. K. Jain, S.R.K.Iyengar and R.K.Jain, "Numerical Methods for Scientific and Engineering Computation", Wiley Eastern Limited, 1993.

[3] M.Gopal," Control Systems ", Tata McGraw Hill @ 2002 pp.314-324

[4] B.N.Datta and Karabi Datta, "On Finding Eigenvalue Distribution of a Matrix in Several Regions of the Complex Plane ", IEEE transactions on Automatic control.Vol-AC-31 No.5 May 1986.

[5] Brain .T. Smith, "Error Bounds for zeros of a polynomial Based upon Gerschgorin theorem ", Journal of the A modulation of Computing Machinery, Vol 17, No.4 October 1970, pp.661-674.

[6] I.J.Nagrath ,M.Gopal , " Control systems Engineering ", Published by New Age International (p) Ltd., Fifth Edition-2007. Reprint-2008.

[7] Yogesh.Vijay Hote, "New Approach of Kharitonov and Gerschgorin theorem in Control systems", A thesis submitted in fulfillment of the requirement in the award of Doctor of Philosophy ", Under the guidance of Prof.J.R.P.Gupta and Prof.Roy Choudhary – Dec-2008.

[8] Y.V.Hote," Dissertation on some interesting Results On the Stability of the systems Matrix a via Gerschgorin theorem ", Submitted to Pune University (India) 1999.

Machining Parameter Optimization of Poly Tetra Fluoro Ethylene (PTFE) Using Genetic Algorithm

M.Sanjeev kumar[1], V.Kaviarasan[2], R.Venkatesan[3]

1 PG Scholar - Product Design and Development
Sona College of Technology, Salem-5
2 Assistant professor, Department of Mechanical Engineering,
Sona College of Technology, Salem-5
3 professor/Head, Department of Mechanical Engineering
Sona College of Technology, Salem-5

ABSTRACT

Poly Tetra Fluoro Ethylene (PTFE) has emerged as an important class of material in aerospace air-conditioning systems, which are increasingly being utilized in recent years. Application of these materials in many areas is due to light weight, corrosion resistant, etc., Surface Roughness (Ra) is an important value for determining the surface quality. Maximum Surface Roughness (Ra) of the tube reduces the air flow pressure, velocity and volumetric air flow rate in air-conditioning systems. Due to high Surface Roughness (Ra) of the air flow tube, the compressor efficiency is reduced and also increases power consumption. The principal machining parameters that control roughness characteristics are cutting speed, feed rate, depth of cut, and type of cutting tools and temperature etc., Genetic algorithm is used for the optimal search of cutting conditions, the chromosomes represent cutting conditions defined according to a sequential scale and is composed by random keys. The present review is focused on the influence of cutting parameters on the surface finish. This result will provide an insight into selecting the optimum machining parameters for machining of PTFE to achieve minimum Surface Roughness (Ra).

Keywords **- PTFE, Surface Roughness, Speed, Feed, Depth of Cut, Genetic Algorithm**

INTRODUCTION

Roughness plays an important role in determining how a real object will interact with its environment. Rough surfaces usually wear more quickly and have higher friction coefficients than smooth surfaces do. Roughness is often a good predictor of the performance of a mechanical component since irregularities in the surface may form nucleation sites for cracks or corrosion. Poly Tetra Fluoro Ethylene

(PTFE) has emerged as an important class of material in aerospace air-conditioning systems, which are increasingly being utilized in recent years. Application of these materials in many areas is due to light weight, corrosion resistant, etc., the term machinability refers to the case with which a material can be machined to an acceptable surface finish. Materials with good machinability require little power to cut, can be cut quickly, easily obtain a good finish, and do not wear the tooling much; such materials are said to be free machining. Machinability can be difficult to predict machining has so many variables. In most cases, the strength and toughness of a material are the primary factors. Strong, tough materials are usually more difficult to machine simply because greater force is required to cut them. Other important factors include the chemical composition, thermal conductivity and microstructure of the material, the cutting tool geometry, and the machining parameters. The machinability can evaluate by different methods. Some of the important methods are Tool life method, Tool forces and power consumption method, Surface finish method and Machinability rating.

In this project, surface roughness is predicted in turning operation using genetic algorithm an optimization technique. Surface finish is an important parameter in this PTFE material because Due to high Surface Roughness (Ra) of the air flow tube, the compressor efficiency is reduced and also increases power consumption. For these reasons, there have been research developments with the objective of optimizing the machining parameters to obtain a good surface finish.

1. Problem Formulation

In machine tools, the finished component is obtained by a number of rough passes and finish passes. The roughing operation is carried out to

machine the part to a size that is slightly larger than the desired size, in preparation for the Finishing cut. The finishing cut is called single-pass contour machining, and is machined along the profile contour.

In this paper, during the turning operation carried out in CNC Lathe under variation of the parameters such as speed, feed, depth of cut in order to minimum surface roughness is predicted.

2. Machining Model

The objective of this model is to minimize the surface roughness. The formula for calculating the above surface roughness is as given by,

$$R_a = -0.309 + 0.675V + 0.870f + 0.175d - 0.234V.f - 0.002f.d - 0.143V.d$$

Finally, by using the above mathematical processes, Surface roughness is obtained.

Where,

V = Cutting Speed (m/min)
f = Feed Rate   (mm/rev)
d = Depth of Cut   (mm)

Machine range:

1. Machine   : CNC Lathe
2. Speed range : 150 – 275 m/min
3. Feed range   : 0.1 – 0.3 mm/rev
4. Depth of cut : 0.5 – 2.5 mm

Outstanding properties of PTFE:

1. Chemical Inertness
2. Non Stick
3. Low Friction
4. Self Lubricating
5. Dielectric Strength
6. Weather Resistance/Non Ageing
7. Insensitive to UV
8. Non Toxic
9. Broad Temperature Range (-200°C to 260°C)
10. Non Flammable
11. Water Absorption = 0!

3. Genetic Algorithm

Genetic algorithm [4, 6, 7] is an adaptive search and optimization algorithm that mimics the principles of natural genetics. GA's are very different from traditional search and optimization methods used in engineering design problems. Because of their simplicity, easy of operations minimum requirements and global perspective, GA's has been successfully used in a wide variety of problem domains. GA work through three operators, namely reproduction, cross over and mutation. In this paper an attempt is made to use of genetic algorithm to minimize the surface roughness by optimizing the depth of cut, feed rate and cutting speeds.

3.1 Steps in the Genetic Algorithm Method

Step 1: Initialization
Randomly generate an initial population of N chromosomes and evaluate the fitness function to a function to be maximized for the encoded version) for each of the chromosomes.

Step 2: Parent Selection
Set if elitism strategy is not used; otherwise. Select with replacement parents from the full population (including the elitist elements). The parents are selected according to their fitness, with those chromosomes having higher fitness value being selected more often.

Step 3: Crossover
For each pair of parents identified in Step 1, perform crossover on the parents at a randomly (perhaps uniformly) chosen splice point (or points if using multi-point crossover) with probability. If no crossover takes place (probability), then form two offspring that are exact copies (clones) of the two parents.

Step 4: Replacement and Mutation
While retaining the best chromosomes from the previous generation, replace the remaining chromosomes with the current population of offspring from Step 2. For the bit-based implementations, mutate the individual bits with probability ; for real coded implementations, use an alternative form of "small" modification (in either case, one has the option of choosing whether to make the elitist chromosomes candidates for mutation).

Step 5: Fitness and End Test
Compute the fitness values for the new population of N chromosomes. Terminate the algorithm if the stopping criterion is met or if the budget of fitness function evaluations is exhausted; else return to Step 1.

Genetic Algorithm Parameters:
1. Population size: 20
2. Chromosome length: 30
3. Selection mode: rank order
4. Cross over: single-site cross over
5. Probability: 0.08
6. Mutation probability: 0.1
7. Fitness: minimum surface roughness

3.2 Working Principle

1. The decision variables X_i are coded in some string structure, binary coded string having zeros and ones are mostly used.

2. The length of the string is usually determined according to the desired solution accuracy. For example, the strings (0000) and (1111) represent the point $(X_1^{(L)}, X_2^{(L)})$ and $(X_1^{(u)}, X_2^{(u)})$, the sub string has the minimum and maximum decoded values.

3. The parameter values are calculated by using the following formula,

$$X = X_i^{(L)} + \frac{X_i^{(U)} - X_i^{(L)}}{2^n - 1} \ (Deocoded\ value)$$

$$(Or)$$

$$x = Min + \left(\frac{Max - Min}{2^n - 1} \right) * (Decoded\ value)$$

3.3 Fitness Function [4]

1. Genetic Algorithm mimics the survival of the fittest principle of nature to make search procedure
2. The fitness function F (x) is first derived from the objective function and used in successive genetic operation
3. For minimization problems, the fitness function is an equivalent maximization problem such that the optimum point remains unchanged.

$$F(X) = \frac{1}{1 + G(X)}$$

3.4 Operation of genetic Algorithm

Genetic Algorithm [4, 6, 7] begins with population of random strings representing design and decision variables thereafter each string is evaluated to find the fitness value.

1. The population is operated by three main operators
    a. Reproduction
    b. Crossover
    c. Mutation
2. The population formed is further evaluated and tested for termination. If the termination criteria is not met, the population is iteratively operated by the above three operators and evaluated.

3. This procedure is continued until the termination criteria are met.

3.5 Genetic Algorithm operators [4, 2]
Reproduction

Reproduction selects good strings in a population and forms a mating pool. The reproduction operator is also called a selection operator. In this work rank order selection is used. A lower ranked string will have a lower fitness value or a higher objective function and vice versa. the probability of selection for each string which is calculated, based on the following formula:

Expected value of probability,

$$= \frac{Min + \square \ (max - min) \ (rank - 1)}{N - 1}$$

Where,   N   = 20
         Min = 0.02
         Max = 0.08

Crossover

In the crossover operator, exchanging information among strings of the mating pool creates new strings. In most crossover operators, two strings picked from the mating pool at random and some portion of the strings are exchanged between the strings.

Mutation

After a crossover is performed, mutation takes place. This is to prevent falling all solutions in population into a Local optimum of solved problem. Mutation changes randomly the new offspring. For binary encoding we can switch a few randomly chosen bits from 1 to 0 or from 0 to 1.

Mutation can then be following:
Before crossover
00011110110001110 | 1100010111011
01101010110011101 | 0100001000100

After crossover
00011110110001110 | 0100001000100
01101010110011101 | 1100010111011

Table 1. Output of parameter values

S.No	Decoded values			String 1	String 2	String 3	Actual speed	Actual feed	Actual depth of cut
1	321	131	244	0101000001	0101000001	0011110100	189.2229	0.138416	1.096285
2	445	231	123	0110111101	0110111101	0001111011	204.3744	0.167742	0.800587
3	323	421	456	0101000011	0101000011	0111001000	189.4673	0.22346	1.61437
4	123	122	268	0001111011	0001111011	0100001100	165.0293	0.135777	1.154936
5	433	432	287	0110110001	0110110001	0100011111	202.9081	0.226686	1.201369
6	66	95	367	0001000010	0001000010	0101101111	158.0645	0.127859	1.396872
7	499	76	287	0111110011	0111110011	0100011111	210.9726	0.122287	1.201369
8	123	118	187	0001111011	0001111011	0010111011	165.0293	0.134604	0.956989
9	403	95	156	0110010011	0110010011	0010011100	199.2424	0.127859	0.881232
10	196	75	99	0011000100	0011000100	0001100011	173.9492	0.121994	0.741935
11	348	372	271	0101011100	0101011100	0100001111	192.522	0.209091	1.162268
12	460	162	269	0111001100	0111001100	0100001101	206.2072	0.147507	1.15738
13	480	423	276	0111100000	0111100000	0100010100	208.651	0.224047	1.174487
14	82	23	313	0001010010	0001010010	0100111001	160.0196	0.106745	1.264907
15	445	345	319	0110111101	0110111101	0100111111	204.3744	0.201173	1.27957
16	234	323	260	0011101010	0011101010	0100000100	178.5924	0.194721	1.135386
17	456	123	342	0111001000	0111001000	0101010110	205.7185	0.13607	1.335777
18	182	456	499	0010110110	0010110110	0111110011	172.2385	0.233724	1.719453
19	427	234	68	0110101011	0110101011	0001000100	202.175	0.168622	0.666178
20	324	234	198	0101000100	0101000100	0011000110	189.5894	0.168622	0.983871

Table 2. After Cross Over

0111110011000100110 1 0111001000
0110111101001110001 1 0101010110
0111100000011010011 1 0100010100
0110111101010101011001 0100111111
0111001000000011111111 0001111011
0111001100001010001 0 0100001101
0111100000011010011 1 0100010100
0101000001101101001 00 0100011111
0001000010000101100 0 0111110011
0001111011000111101 0 0100001100
0111100000011010011 1 0100010100
0010110110011100111 1 0101101111
0001111011000111001 0 0001000100
0110010011000101101 1 0101010110
0110101011001110111 0 0010111011
0101000001001000011 0 0010111011
0101011100010111010 0 0100001111
0001000010000101111 1 0101101111
0001111011000111001 1 0011110100
0111001000000011111111 0010011100

4. GA Procedure[4]

Step1:

Choose a coding to represent problem parameter, a selection operator, a crossover operator and a mutation operator. Choose population size N, crossover probability p_c, and mutation probability p_m. Initialize a random population of strings of size 10. Set iteration t=0.

Step 2: Evaluate each string in the population.

Step 3: If it > it_{max} (or) other termination criteria is satisfied, terminate.

Step 4: Perform reproduction on the population.

Step 5: Perform crossover on the random pairs of strings.

Step 6: Perform bit wise mutation.

Step 7: Evaluate strings in the new population. Set it = it +1 and go to step 3.

Table 3.Output of Genetic Algorithm

S.No	Surface roughness	fitness	Sorted fitness	Rank	Probability of selection	Cumulative probability	Random number	Selected rank
1	91.93534	0.01076	0.009076	1	0.02	0.02	0.237122	7
2	106.5099	0.009301	0.009301	2	0.023158	0.043158	0.055939	2
3	74.41097	0.013261	0.009581	3	0.026316	0.069474	0.515228	13
4	78.90687	0.012515	0.009888	4	0.029474	0.098947	0.678711	15
5	91.43897	0.010818	0.010105	5	0.032632	0.131579	0.783447	17
6	70.43693	0.013998	0.010422	6	0.035789	0.167368	0.491699	12
7	100.1326	0.009888	0.010535	7	0.038947	0.206316	0.561035	13
8	83.58795	0.011822	0.010545	8	0.042105	0.248421	0.083893	3
9	103.376	0.009581	0.010632	9	0.045263	0.293684	0.17981	6
10	93.92137	0.010535	0.01076	10	0.048421	0.342105	0.10199	4
11	88.61061	0.011159	0.010818	11	0.051579	0.393684	0.517456	13
12	97.96541	0.010105	0.010867	12	0.054737	0.448421	0.85709	18
13	94.94815	0.010422	0.011159	13	0.057895	0.506316	0.272369	8
14	75.07653	0.013145	0.011822	14	0.061053	0.567368	0.340149	9
15	91.02521	0.010867	0.011838	15	0.064211	0.631579	0.938477	19
16	83.47469	0.011838	0.012515	16	0.067368	0.698947	0.032898	1
17	93.05703	0.010632	0.013145	17	0.070526	0.769474	0.436371	11
18	64.68515	0.015224	0.013261	18	0.073684	0.843158	0.198944	6
19	109.185	0.009076	0.013998	19	0.076842	0.92	0.266113	8
20	93.82769	0.010545	0.0159	20	0.08	1	0.774445	17

Fig 1. Flow chart
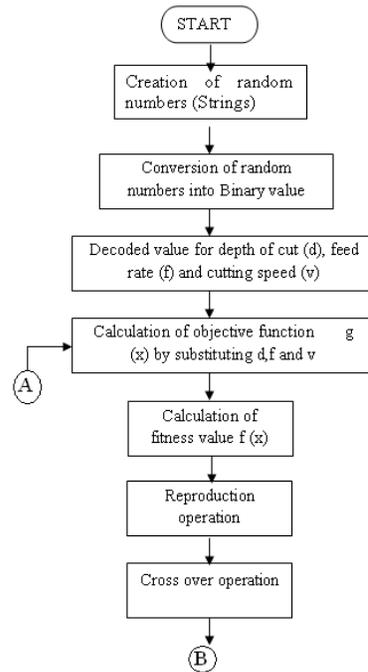
Objective function:

The objective of this model is to minimize the Surface roughness.

The formula for calculating the Surface roughness is as given by,

$$R_a = -0.309 + 0.675V_c + 0.870f + 0.175d - 0.234V_c.f - 0.002f.d - 0.143Vcd$$

Finally, by using the above mathematical processes, the Surface roughness is obtained.

Where, V   = Cutting Speed   (m/min)
        f   = Feed Rate   (mm/rev)
        d   = Depth of Cut   (mm)
        Ra   = Surface roughness   (µm)
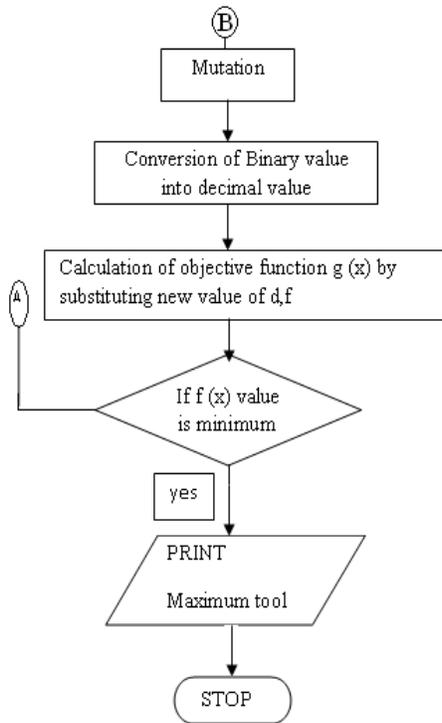
Graphical output of genetic algorithm
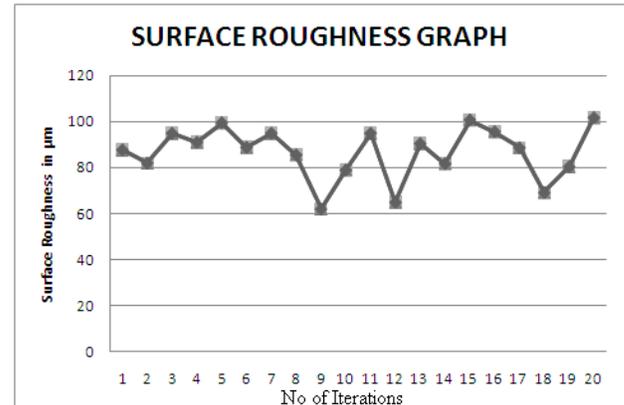
Table 4. Output after iteration

S.No	Decoded values			Children speed (m/min)	Children feed (mm/rev)	Children DOC (mm)	Surface roughness (µm)	Fitness
1	499	77	456	210.9726	0.122581	1.61437	87.73072	0.01127
2	445	227	470	204.3744	0.166569	1.648583	81.92994	0.012058
3	480	423	276	208.651	0.224047	1.174487	94.94815	0.010422
4	445	345	319	204.3744	0.201173	1.27957	91.02521	0.010867
5	456	127	251	205.7185	0.137243	1.113392	99.50479	0.00995
6	460	162	397	206.2072	0.147507	1.470186	88.79617	0.011136
7	480	423	276	208.651	0.224047	1.174487	94.94815	0.010422
8	323	420	287	189.4673	0.223167	1.201369	85.5414	0.011555
9	66	216	499	158.0645	0.164433	1.719453	61.92024	0.015893
10	123	122	268	165.0293	0.135777	1.154936	78.90687	0.012515
11	480	423	276	208.651	0.224047	1.174487	94.94815	0.010422
12	234	323	260	178.5924	0.194721	1.135386	83.47469	0.011838
13	123	114	68	165.0293	0.133431	0.666178	90.44433	0.010936
14	403	91	470	199.2424	0.126686	1.648583	81.70061	0.012092
15	427	238	187	202.175	0.169795	0.956989	100.7736	0.009826
16	321	134	187	189.2229	0.139296	0.956989	95.64201	0.010347
17	348	372	271	192.522	0.209091	1.162268	88.61061	0.011159
18	66	223	367	158.0645	0.165396	1.396872	69.08111	0.014269
19	123	115	244	165.0293	0.133724	1.096285	80.35823	0.012291
20	456	127	220	205.7185	0.137243	1.037634	101.7202	0.009735

5. RESULT AND DISCUSSION

The objective function is the minimization of surface roughness by varying feed, speed, depth of cut. In this work, the optimum surface roughness is obtained by using genetic algorithm at the 9th generation. The optimum value of surface roughness is **61.92024**μm. The corresponding speed is **158.0645** m/min, feed is **0.164433**mm/rev and depth of cut is **1.719453** mm. These are the best parameters obtained to achieve minimum surface roughness in machining PTFE tubes to enhance the airflow in the airplanes and aircraft air-conditioning systems.

6. CONCLUSION

GA's are derivative-free calculations and therefore, are neither bound to assumptions regarding continuity, nor limited by required prerequisites. As Goldberg stated, GAs are blind. They can handle any kind of objective function and any kind of constraints (e.g., linear or nonlinear) defined on discrete, continuous or mixed search spaces. In addition, as stated earlier, they are robust in producing near-optimal solutions, with a high degree of probability to obtain the global Optimum. A genetic algorithm was proposed for optimizing the machining parameters. The main advantage of this approach is that it can be used for any objective function, which was most clearly demonstrated in this example, where the objective function was the minimization of surface roughness. In this approach three constraints namely feed, speed and depth of cut are considered for minimizing surface roughness in PTFE tubes of air-conditioning systems. There are many other constraints that affect surface roughness, which can be solved by using multi objective genetic algorithm in the future.

References:

Journal Papers:

[1]   Ramon Quiza Sardinas, Marcelino Rivas Santana and Eleno Alfonso Brindis, Genetic algorithm-based multi-objective optimization of cutting parameters in turning processes, Engineering Applications of Artificial Intelligence, 19, 2006, 127 – 133

[2]   R. Saravanan, P. Asokan and K Vijayakumar, Machining Parameters Optimisation for Turning Cylindrical Stock into a Continuous Finished Profile Using Genetic Algorithm (GA) and Simulated Annealing (SA), *International Journal of Advanced Manufacturing technology,* l, 2003, 21:1–9

[3]   M. C. Chen and C. T. Su, Optimisation of machining conditions for turning cylindrical stocks into continuous finished profiles, *International Journal of Production Research*, *36(8), pp. 2115–2130*, 1998.

[4]   Narayanasamy, R., Venkatesan, R. and Ponalagusamy, R., Extrusion Die Profile And Extrusion Pressure Optimization Using Genetic Algorithm, *International Journal of Institution Engineers, Singapore, Vol.45, pp 1-13*, 2005.

[5]   B. Gopalakrishnan et al., Machine parameter selection for turning with constraints: an analytical approach based on geometric programming*, International Journal of Production Research, 29(9),pp. 1897–1908,* 1991

Books:

[6]   Kalyanmoy Deb, *Optimizations for engineering design – Algorithm and example*s (Prentice-Hall of India, New Delhi, 1996).

[7]   S.Rajasekaran,G.A.Vijayalakshmi pai, *Neural Networks, Fuzzy Logic and Genetic Algorithms Synthesis and Applications*(PHI Learning pvt ltd,new delhi, 2003)

[8]   D.E. Goldberg, *Genetic Algorithm in search optimization and Machine Learning(* Addison Wesley, MA 1989).

Proceedings Papers:

[9]   Maria E. Requena-Pérez, Antonio Albero-Ortiz, Juan Monzó-Cabrera, and Alejandro Díaz-Morcillo, Combined Use of Genetic Algorithms and Gradient Descent Optmization Methods for Accurate Inverse Permittivity Measurement, *IEEE transactions on Microwave Theory and Techniques, vol. 54, no. 2, February 2006.*

Dynamic Test Generation to Find Bugs in Web Application

C .SathyaPriya[1] and S.Thiruvenkatasamy [2]

[1] Department of IT, Shree Venkateshwara Hi-Tech Engineering College, Gobi, Tamilnadu, India.
[2] Department of CSE, Shree Venkateshwara Hi-Tech Engineering College, Gobi, Tamilnadu, India.

Abstract- **Current tools for webpage validation cannot handle the dynamically generated pages that are ubiquitous on today's Internet. There are no approaches which can detect the failures in dynamically generated web applications without executing them for several times. The Developer of the dynamic web applications has to check for failures only after executing the application in web server. If there are many failures occur, he has to correct them and again executes the application several times We present a dynamic test generation technique for the domain of dynamic Web applications. The technique utilizes both combined concrete and symbolic execution. This paper extends dynamic test generation to the domain of web applications that dynamically create web (HTML) pages during execution, which are typically presented to the user in a browser. We are proposing such approach which can detect the execution and HTML failures in dynamic web applications before they are going to be executed.**

*Keywords-***Software testing, Web applications, HTML.**

1. INTRODUCTION

Web developers widely recognize the importance of creating legal HTML. Many websites are checked using HTML validators. Recently, several approaches have been proposed to address Web application testing [1, 2, 4, 5, 6]. Most of the approaches focus on testing the architectures of Web Applications and they are not applicable for testing malformed dynamic web pages and web script crashes. These seriously impact the usability of web applications. There are no approaches which can detect the failures in dynamically generated web applications without executing them for several times. If there are many failures occur, he has to correct them and again executes the application several times. The present tools are capable to test either the java code or the validating HTML content of the web page after the web application executed. Thus, we present an approach which can detect the execution and HTML failures in dynamic web applications that are developed using JSP pages, a very popular server-side script language for developing Web applications with Java technology.

There are two general approaches to finding faults in web applications: static analysis and dynamic analysis (testing).

In the context of Web applications, static approaches have limited potential because 1) Web applications are often written in dynamic scripting languages that enable on-the-fly creation of code, and 2) control in a Web application typically flows via the generated HTML text (e.g., buttons and menus that require user interaction to execute), rather than solely via the analyzed code. Both of these issues pose significant challenges to approaches based on static analysis. Testing of dynamic Web applications is also challenging because the input space is large and applications typically require multiple user interactions.

The state of the practice in validation for Web-standard compliance of real Web applications involves the use of programs such as HTML Kit5 that validate each generated page, but require manual generation of inputs that lead to displaying different pages. We know of no automated tool that automatically generates inputs that exercise different control-flow paths in a Web application, and validates the dynamically generated HTML pages that the Web application generates when those paths are executed.

This paper presents an automated technique for finding failures in HTML-generating web applications. Our technique is based on dynamic test generation, using combined concrete and symbolic (concolic) execution, and constraint solving .We created a tool, Apollo, that implements our technique in the context of the publicly available PHP

interpreter. Apollo first executes the Web application under test with an empty input. During each execution, Apollo monitors the program to record path constraints that reflect how input values affect control flow. Additionally, for each execution, Apollo determines whether execution failures or HTML failures occur (for HTML failures, an HTML validator is used as an oracle). Apollo automatically and iteratively creates new inputs using the recorded path constraints to createinputs that exercise different control flow.

2. RELATED WORK

An earlier version of this paper was presented at ISSTA '08 [2]. The Apollo tool presented there did not handle the problem of automatically simulating user interactions in Web applications. Instead, it relied on a manual transformation of the program under test to enable the exploration of a few selected user inputs. The current paper also extends [2] by providing a more extensive evaluation, which includes two new large Web applications, and by presenting a detailed classification of the faults found by Apollo. In addition, the Apollo tool presented in [2] did not yet support Web server integration. In the remainder of this section, we discuss three categories of related work:
1) combined concrete and symbolic execution,
2) techniques for input minimization, and
3) testing of Web applications. Apollo first executes the Web application under test with an empty input.

During each execution, Apollo monitors the program to record path constraints that reflect how input values affect control flow. Additionally, for each execution, Apollo determines whether execution failures or HTML failures occur (for HTML failures, an HTML validator is used as an oracle). Apollo automatically and iteratively creates new inputs using the recorded path constraints to create inputs that exercise different control flow. Most previous approaches for concolic execution only detect "standard errors" such as crashes and assertion failures.

2.1 Combined Concrete and Symbolic Execution

DART[5] is a tool for finding combinations of input values and environment settings for C programs that trigger errors such as assertion failures, crashes, and no termination. DART combines random test generation with symbolic reasoning to keep track of constraints for executed control flow paths. A constraint solver directs subsequent executions toward uncovered branches. Experimental results indicate that DART is highly effective at finding large numbers of faults in several C applications and frameworks, including important and previously unknown security vulnerabilities. CUTE is a variation (called concolic testing) on the DART approach. The authors of CUTE introduce a notion of approximate pointer constraints to enable reasoning over memory graphs and handle programs that use pointer arithmetic. Subsequent work extends the original approach of combining concrete and symbolic executions to accomplish two primary goals:

1) Improving scalability
2)Improving execution coverage and fault detection capability.

Godefroid proposed a compositional approach to improve the scalability of DART. In this approach, summaries of lower level functions are computed dynamically when these functions are first encountered.

3. LIMITATIONS

Simulating user inputs based on locally executed Java- Script. The HTML output of a PHP script might contain buttons and arbitrary snippets of JavaScript code that are executed when the user presses the corresponding button. The actions that the JavaScript interpreter might perform are currently not analyzed by Apollo. For instance, the Java- Script code might pass specific arguments to the PHP script. As a result, Apollo might report false positives. For example, Apollo might report a false positive if Apollo decides to execute a PHP script as a result of simulating a user pressing a button that is not visible. Apollo might also report a false positive if it attempts to set an input parameter that would have been set by the JavaScript code. In our experiments, Apollo did not report any false positives.

Apollo has limited tracking of input parameters through PHP native methods. PHP native methods are implemented in C, which make it difficult to automatically track how input parameters are transformed into output parameters. We have modified the PHP interpreter to track parameters across a very small subset of the PHP native methods. Similarly to we plan to create an external language to model the dependencies between

inputs and outputs for native methods to increase Apollo line coverage when native methods are executed.

Apollo does not track input parameters through the database. Thus, Apollo might not be able to explore call sequences in which subsequent calls depend on specific values of input parameters stored in the database by earlier calls. It is possible to extend Apollo to track input parameters through the database in the same way Emmi et al. extended concolic testing to database applications.

In theory, Apollo might be unable to cover certain parts of an application because the constraints that use are fairly simple. However, from what we have observed in our experiments so far, PHPWeb applications do not seem to manipulate their input parameters in complex ways, and the very simple constraints that can be solved with Choco have been adequate for our purposes. We have only observed a very few isolated cases where the solving of more complex constraints would have helped.

3. IMPLEMENTATION

3.1 Web Application

HTML and JSP are widely used for implementing web applications, in part due to its rich library support for network interaction, HTTP processing, and database access. The input to a HTML program is a map from strings to strings. Each key is a parameter that the program can read, write, or check if it is set. The string value corresponding to a key may be interpreted as a numerical value if appropriate. The output of a JSP web application is an HTML document that can be presented in a web browser. JSP is object-oriented, in the sense that it has classes, interfaces, and dynamically dispatched methods with syntax and semantics.

3.2 Finding failures in HTML web Application

This technique targets two types of failures that can be automatically identified during the execution of web applications. First, execution failures may be caused by a missing included file, an incorrect MySQL query, or an uncaught exception. Such failures are easily identified as the interpreter generates an error message and halts execution. Second, HTML failures involve situations in which the generated HTML page is not syntactically correct according to an HTML validator. The basic idea

behind the technique is to execute an application on some initial input (e.g., an arbitrarily or randomlychosen input), and then on additional inputs obtained by solving constraints derived from exercised control flow paths.

3.3 Path constraint minimization

The failure detection algorithm returns bug reports. Each bug report contains a set of path constraints, and a set of inputs exposing the failure. Previous dynamic test generation tools presented the whole input (i.e., many _input Parameter, value_ pairs) to the user without an indication of the subset of the input responsible for the failure. As a postmortem phase, this minimization algorithm attempts to find a shorter path constraint for a given bug report.

The set of all such required conjuncts determines the minimized path constraint. From the minimized path constraint, the algorithm produces a concrete input that exposes the failure. This approach for minimization is similar in spirit to delta debugging, a well-known input minimization technique. The current minimization algorithm also differs from delta debugging in that it does not rely on efficient binary search like techniques to identify redundant components of path constraints.

3.4 Concrete and symbolic execution

Symbolic execution is a form of program analysis that uses symbolic values instead of actual data as inputs and symbolic expressions to represent the values of program variables. As a result, the outputs computed by a program are expressed as a function of the symbolic inputs. The state of a symbolically executed program includes the (symbolic) values of program variables, a path condition (PC), and a program counter.

The path condition is a Boolean formula over the symbolic inputs, encoding the constraints which the inputs must satisfy in order for an execution to follow the particular associated path. The paths followed during the symbolic execution of a program are characterized by a symbolic execution tree.

3.5 Implementation

Own tool is created that implements the technique for HTML. This own tool consists of three major components, Executor, Bug Finder, and Input Generator . This section first provides a high-level overview of the components and then discusses the pragmatics of the

implementation. The inputs to own tool are the program under test and an initial value for the environment. The initial environment usually consists of a database populated with some values, and user-supplied information about username/password pairs to be used for database authentication.

4. TESTING OF WEB APPLICATIONS

Existing techniques for fault detection in Web applications focus on output correctness and security. Minamide uses static string analysis and language transducers to model PHP string operations to generate potential HTML output—represented by a context-free grammar—from the Web application. This method can be used to generate HTML document instances of the resulting grammar and to validate them using an existing HTML validator.

Benedikt et al. present a tool, VeriWeb, for automatically testing dynamic webpages. They use a model checker to systematically explore all paths (up to a certain bound) of user navigatable components in a website.

When the exploration encounters HTML forms, VeriWeb uses SmartProfiles. SmartProfiles are user-specified attributevalue pairs that are used to automatically populate forms and supply values that should be provided as inputs. Although VeriWeb can automatically fill in the forms, the human tester needs to prepopulate the user profiles with values that a user would provide. Similarly, the WAVES tool by Huang et al. performs automatic form completion by using a textual analysis to associate "topics" with input values that occur in HTML forms, in combination with a self-learning knowledge base that associates values with topics.

In contrast, Apollo automatically discovers input values based on the examination of branch conditions on execution paths. Benedikt et al. do not report any faults found, while we report 673. Halfond and Orso [use static analysis of the serverside implementation logic to extract a Web application's interface, i.e., the set of input parameters and their potential values. (Halfond et al. later extended that approach to include dynamic analysis and specialized constraint solving.) They implemented their technique for Web applications written in Java. They obtained better code coverage with test cases based on the interface extracted using their technique as compared to the test cases based on the interface extracted using a conventional Web crawler.

4. CONCLUSIONS AND FUTURE WORK

A technique is presented for finding faults in Web applications that is based on combined concrete and symbolic execution. We performed our own tool and compromised the execution failures and as well as the HTML failures. The work is novel in several respects.

First, the technique not only detects runtime errors but also uses an HTML validator as an oracle to determine situations where malformed HTML is created. *Second*, we address a number of PHP-specific issues, such as the simulation of interactive user input that occurs when user-interface elements on generated HTML pages are activated, resulting in the execution of additional PHP scripts. *Third,* we perform an automated analysis to minimize the size of failure-inducing inputs.

Our approach is the first one to detect failures of a dynamic web application before execution and also without deploying it for several times in order to test the web applications. Though we limited our approach due to the complex nature of Java based dynamic web applications, our technique works fine to find the failures of both execution and HTML.

In future, we try to work on the limitations and to detect more number of failures in both cases and also improve the usability of our technique.

5. REFERENCE

[1]. Shay Artzi, Adam Kie _zun, Julian Dolby, Frank Tip, Danny Dig, Amit Paradkar, Senior Member, IEEE, and Michael D. Ernst, "Finding Bugs in Web Applications Using Dynamic Test Generation and Explicit-State Model Checking" VOL. 36, NO. 4, JULY/AUGUST 2010.

[2]. J. Clause and A. Orso, "Penumbra: Automatically Identifying Failure-Relevant Inputs Using Dynamic Tainting," Proc. Int'l Symp. Software Testing and Analysis, 2009.

[3]. W.G. Halfond, S. Anand, and A. Orso, "Precise Interface Identification to Improve Testing and Analysis of Web Applications," roc. nt'l Symp. Software Testing and Analysis, 2009.

[4]. S. Sinha, H. Shah, C. Go¨rg, S. Jiang, and M. Kim, "Fault Localization and Repair for Java Runtime Exceptions," Proc. Int'l Symp. Software Testing and Analysis, 2009.

[5] P. Godefroid, N. Klarlund, and K. Sen, "DART: Directed Automated Random Testing," Proc. ACM SIGPLAN Conf. Programming Language Design and Implementation, pp. 213-223, 2005.

[6] Y.Minanide, "Static Approximation of Dynamically Generated Web Pages", Proc. Int'1 Conf. World Wide Web 2005.

[7] P. Godefroid, A. Kiezun, and M. Y. Levin. Grammar-based whitebox fuzzing. *In PLDI, 2008*.

BIOGRAPHY

C.Sathyapriya, received her Post Graduate Degree in Master of Engineering in Computer Science, from Velalar College Of Engineering and Technology, Erode, Tamil Nadu, India. Currently she is working as Assistant Professor in the Department of Information Technology in Shree Venkateshwara Hi-Tech Engineering College, Gobichettipalayam, Tamil Nadu, India. She had presented 10 papers in Various National Conference and also she presented 1 International Conference on "A Test Generation Method to Find Errors in HTML Language" in VIT University, Vellore, Tamil Nadu, India on 21st April 2011. Her interest includes Software Engineering, Computer Networks and Data Mining. She has Published 1 paper in National journal.

S.Thiruvenkatasamy, received his Post Graduate Degree in Master of Engineering in Computer Science, from Karpagam University, Coimbatore, Tamil Nadu, India. Currently he is working as Assistant Professor in the Department of Computer Science and Engineering in Shree Venkateshwara Hi-Tech Engineering College, Gobichettipalayam, Tamil Nadu, India. He published a book "An Excellent Guide for Visual C#.Net" and had presented 6 Papers in Various National Conferences. His interest includes Data mining, computer networks and Network security. He has Published 1 paper in National journal.

A study on the application of Elliptic - Curve cryptography in implementing smart cards

Yasir Ahmad

Manav Bharti University, India

ABSTRACT

Elliptic curve cryptography is one of the emerging techniques that stand as an alternative for conventional public key cryptography. Elliptic curve cryptography has several applications of which smart cards are also one among them. A smart card is nothing but a single chip that contains microprocessor components. Smart cards are mainly used for secured sign-on in big organizations. The security feature of smart card is provided by elliptic curve cryptography. Elliptic curve cryptography for smart cards can be implemented through several ways. Of them implementation using Galois Field is one of the very famous techniques. This essay discusses in detail how elliptic curve cryptography is implemented in smart cards using a concept called Galois finite field.

Keywords — **Elliptic curve cryptography, Cryptography.**

1. AN OVERVIEW OF ELLIPTIC CURVE THEORY:

Elliptic curves are referred so because they are explained by triple equations, similar to those used in the calculations of ellipsis [1]. The elliptic curve equation general form is:

$$b2 + cab + db = a3 + ea2 + fa + g$$

The below figure shows an example of ECC curve:
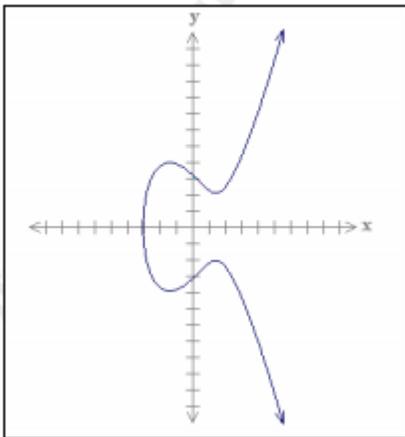
Fig 1: ECC Example
Source: sans.org

1.1 What is Elliptic Curve Cryptography?

Elliptic curve cryptography was introduced by Neal Kolbitz and V Miller in 1985. Elliptic Curve Cryptography proposed as an alternative to established public key systems such as RSA and has recently achieved lot of attention in academia and industry [2]. The major cause for the elliptic curve cryptography attractiveness is the fact that there is no sub exponential algorithm known to solve the discrete logarithm issue on an appropriately selected elliptic curve. This means that importantly smaller parameters can be used in Elliptic Curve Cryptography than in other competitive systems such as DSA and RSA but with similar security levels. Some advantages of having little key sizes include reductions and quicker computations in storage space, bandwidth and processing power. This makes Elliptic curve cryptography for constrained building of elliptic curve cryptography. Such as Personal digital assistants, pagers, smart cards and cellular phones. On the other hand the elliptic curve cryptography implementation needs many options such as the kind of the underlying finite field, algorithms for establishing the finite field arithmetic and so on.

Contrary to that Tilborg and Jajodia [3] defined that elliptic curve cryptography enhances the analysis and configuration of public key cryptographic schemes that can be established using elliptic curves. The elliptic curve scheme analogues based on the discrete logarithm issue where the underlying group is the collection of points on an elliptic curve defined over a finite field.

Stavroulakis and Stamp [4] described that elliptic curve cryptography enhances using the group of points on an elliptic curve as the underlying number system for public key cryptography. There are two major causes for using elliptic curves as a basis for public key cryptosystems. The first reasons are that the elliptic curve based cryptosystems exists to offer better security than traditional cryptosystems for a given key size. One can take benefit of this fact is to develop security or to develop performance by lowering down the size of the key while keeping common security. The second cause is that the additional framework on an elliptic curve can be destructed to build cryptosystems with interesting features which are impossible or critical to gain in any other way.

Elliptic curves are algebraic structures that form a basic class of cryptographic primitives which depend on a mathematical hard issue. The elliptic curve discrete algorithms

problem is based on the intractability of deriving a huge scalar after its multiplications with a given point on an elliptic curve Yalcin [5].

According to Zheng and Lionel [6] an alternative to RSA elliptic curve cryptography is another approach to public key cryptography. Elliptic curve cryptography is based on the property of elliptic curve in algebraic geometrics. The elliptic curve cryptography permits one to select a secret number as a private key which is then used to select a point on a non secret elliptic curve. A nice property of an elliptic curve is that it enhances both parties to compute a secret key solely based on its private key and other's public key.

2. WHAT IS SMART CARD?

Smart cards are perhaps some of the most vastly used electronic components in use nowadays Mayes and Markantonakis, [7] define that. The below figure shows the physical appearance of smart card:
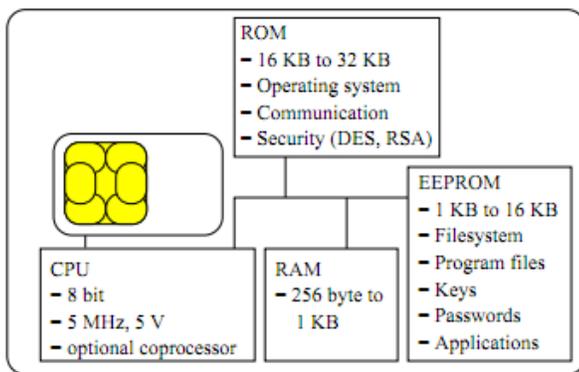

Fig 2: Physical Appearance of Smart Card
Source: Chen (2000), Java Card technology for Smart Cards: architecture and programmer's guide, Sun Microsystems Inc., USA, p 12

Because they have determined to be little and often concealed, smart cards have carried on their necessary works unnoticed largely but this situation is changing today. The high profile use of smart cards for IDs, credit cards, e-tickets and passports means that the smart card has now emerged as a common utility today. A smart card can be used in an automated electronic transaction. It is not easily copied or forged and it is used mainly to add security. Smart cards can also store data protectively and they can run or host a range of security functions and algorithms.

Contrary to that Cranor and Garfinkel [8] defined that smart cards are praised always for their usability. They are mobile and they can be used in several applications and carry lesser administrative costs than systems based on several user name or passwords. On the other hand smart cards are also criticized for their less acceptance of market. Several people use this smart card added choice of security because readers and smart cards are not deployed vastly. However alternative form factors to the familiar plastic smart card are arousing, proponents of these technologies claims that they overcome the smart card limitations.

The smart card is a component which is able to store data and run commands. It is a single chip microcomputer with a size of 25 mm at most. This microcomputer is placed on a plastic card of the size of a standard credit card. Plastic cards have a long tradition. The smart card is a protective and tamper resistant component. The data stored on the card can be protected with a secret which is shared between the smart card and the cardholder. Only the person knowing the secret can use the card and the information stored on it. With the ability to execute commands and programs the smart card became able to decrypt and encrypt information argues, Hansmann [9].

Contrary to that Chen [10] defined that a smart card processes and stores information through the electronic circuits fixed in silicon in the plastic substrate of its body. A smart card is a tamper resistant and portable computer. Unlike magnetic stripe cards, smart cards carry both information and processing power. Therefore they do not need access to remote databases at the transaction time.

According to American Heritage Dictionary [11] a smart card is a small plastic card containing a computer chip. Several smart cards consist of memory chips to store data but several chips also contain microprocessors that can process data. Smart cards are part of systems that enhance cardholders to buy services and goods, enter prohibited areas, links to cell phone networks or operate other operations that needs the processing and storage of recognizing information. SIM cards are a famous kind of smart card.

Beiske, Lee, Yim and Yu [12] have defined that smart card is a credit card size plastic card which consists of magnetic data or microchip area. When this chip consists of monetary information which can be used for later transactions these smart cards belong to the group of electronic cash. Thus at present the electronic cash and smartcards are separated into online and offline applications. Some of the smartcard issues are that the smart cards are virtual and real stores accept them and are secure, efficient, speedy, paperless and intuitive. Smart card also supports several industries from banking to health care.

3. TYPES OF SMART CARDS

Smart cards fall into different groups. They can be categorized into microprocessor cards, memory cards, contactless cards and contact cards based on the variations in the access mechanism of cards. The types of smart cards are explained below:

3.1 Memory Cards:

The below figure shows the architecture of a memory card in block diagram form:
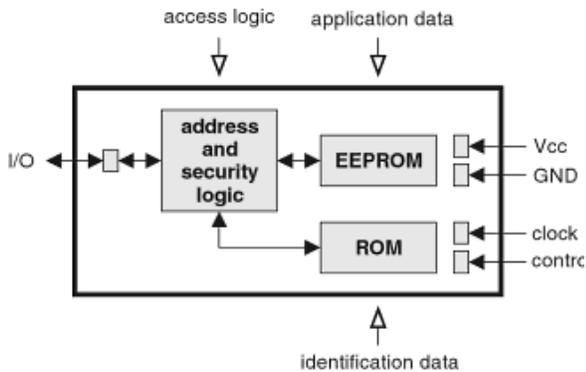
Fig 3: Memory card architecture
Source: Silabs.com

According to Ranki and Effing [13] the required data by the application is stored in nonvolatile memory which is usually referred to as EEPROM. The memory access is controlled by the security logic which in the easiest case contains only of write security or erases security for the memory or specific memory regions. However there are also memory chips with more complex protective logic that can also perform easy encryption. The data is transferred to and from the card through a serial interface. The memory cards functionality is optimized usually for a specific application. Although this severely prohibits the cards flexibility it makes them quite costly. For memory cards typical applications are prepaid telephone cards and easy health insurance cards.

3.2 Microprocessor cards:

Microprocessor cards are sometimes known as microcontroller cards which consist of a microcontroller usually with mask Read Only Memory (ROM) and Electrically Erasable Programmable ROM (EEPROM) for personalization. Nowadays the microcontroller cards can have a technical capability to carry out several functions that are expected of a personal computer. These cards are assumed as truly smart and their processing ability that enhances them to be active and able to react and process to data in a given situation. The ability to perform independent calculations and to store several microprocessor cards applications means that they are suited well for application in transport, banking and some multi-application loyalty systems [14].

3.3 Contact Cards:

The contact cards will be used as replacements for magnetic stripe cards mainly in access control and financial applications. Most of these contact cards will make their interface with the outside globe through a set of 6 to 8 contacts as defined in ISO

7816 part 1. The contact cards are themselves an important point of weakness in a smart card system: 1) The leads from the microcircuit to the contacts are of importance very thin and can become or break detached when the card is stressed or otherwise bent; 2) the contacts can become worn through damaged or excessive use by a defective reader or in a pocket; 3) they denote an obvious initiating point for any attack; and 4) In the reader the contact set is a mechanical component which can break or be damaged either maliciously or accidently [15].

3.4 Contactless cards:

Jurgensen and Guthery [16] described that a contactless card has an ICC embedded within the card. However it makes use of an electromagnetic signal to facilitate communication between the reader and the card. With these cards the important power to run the chip on the card id transformed at microwave frequencies from the reader into the card. The separation permitted between the card and the reader is quite small on the order of a few millimeters. However this card provides a higher ease of use than cards that must be inserted into a reader. This ease of use can be mitigated by other factors.

4. USE OF ELLIPTIC CURVE CRYPTOGRAPHY IN SMART CARDS

Tipton and Krause [17] describe that ECC is suited ideally for implementation in smart cards for several reasons:

- **Scalability:** As the applications of smart card needs stronger and stronger security with big keys, Elliptic curve cryptography can continue to offer the security with proportionately lesser additional system resources. This means that with elliptic curve cryptography smart cards are capable of offering higher security levels without developing their prices.

- **Shorter transmission times and less memory:** The elliptic curve discrete logarithm problem algorithm strength means that strong security is gained with proportionately certificate sizes and smaller key. The smaller size of key in turn means that small memory is needed to store certificates and keys and that less data must be passed between the application and the card so transmission times are shorter.

- **No coprocessor:** The elliptic curve cryptography reduced processing times also make it separate for the platform of smart card. Other public key systems involve many computation that a dedicated hardware component referred to as crypto coprocessor is needed. The crypto coprocessors not only take up huge amount of space on the card but

they also higher the price of the chip by about 20 to 30% which transforms to an increase of about $3 to #5 on the cost of each card. With elliptic curve cryptography the algorithm can be implemented in available Read Only Memory so no extra hardware is needed operate fast and strong functions of security.

- **On card key generation:** As described above the private key in a public key pair must be kept secret. To prevent the transaction truly from being refuted the private key must be inaccessible wholly to all parties except the entity to which it belongs. In applications using the other kinds of public key systems presently in use cards are personalized in a protective environment to meet this need. Because of the complexity of the computation needed generating keys on the card is typically impractical and inefficient.

With Elliptic Curve Cryptography the time required to produce a key pair is so small that even a component with a very limited computing smart card power can produce the key pair offered a better random number generator is possible. This means that the process of card personalization can be streamlined for applications in which no repudiation is necessary.

5. ELLIPTIC CURVE CRYPTOGRAPHY IMPLEMENTATION IN SMART CARDS

In general elliptic curve cryptography is implemented in smart cards by the use of a concept called finite field. Finite field is nothing but a set of elements which have a finite order. Galois Field represented by GF is a finite field whose order is in general a prime number, denoted as GF(m) or power of prime number denoted by GF (2m). The complexity of the arithmetic of the elliptic curve depends upon the finite field in which the elliptic curve is applied. GF(2m) is one of the most popular methods of implementing elliptic curve cryptography. The smart cards can be implemented using GF (2m). With GF (2m) a smart card is less costly because a coprocessor is not required. GF (2m) is referred to as a binary finite field or a two field characteristic. It can be looked as dimension's vector space k over the field GF (2m) that contains of 0 and 1 element. To describe in detail, there occur m elements (y0, y1, y2 . . . , ym-1) in GF (2m) such that every element y ε GF (2m) can be written distinctly in the form:

$$y = b0 \ y0 + b1y1 \ . \ . \ . + bm\text{-}1ym\text{-}1$$
$$\text{where bi } \varepsilon \text{ GF (2)}$$

Such a set { y0, y1, y2 . . . , ym-1} is referred to as GF (2m) basis over GF (2). When such a basis is given a field element y can be denoted as a bit string (b0, b1 . . . bm-1). Performing

extra field elements can be gained simply by XOR-ing bit-wise which are the elements vector representations. The rule of multiplication relies on the chosen basis. GF (2m) over GF (2) has several varied bases. Some bases may lead to several efficient arithmetic implementations in GF (2m) than other bases. The most famous 2 used bases are the normal and polynomial bases. In single representation of basis the elements can be transformed efficiently to other basis representation elements by using proper interoperability and change-of-basis matrix, between systems using 2 varied field types of representation can be gained easily. The equation of elliptic curve over GF (2m) is:

$$a2 + ya = a3 + by2 + c$$
$$\text{where y, a, b, c } \varepsilon \text{ GF (2m) and c} \neq 0$$

The sum of 2 varied points on elliptic curve is evaluated as shown below:

$$(a1, b1) + (a2, b2) = (a3, b3); \text{ where a1} \neq a2$$
$$\lambda = (b2 + b1)/ (a2 + a1)$$
$$a3 = \lambda \ 2 + \lambda + a1 + a2 + c$$
$$b3 = \lambda \ (a1 + a3) + a3 + b1$$

On the elliptic curve the doubling a point is evaluated as shown below:

$$(a1, b1) + (a1, b1) = (a3, b3); \text{ where a1} \neq 0$$
$$\lambda = a1 + (b1) / (a1)$$
$$a3 = \lambda \ 2 + \lambda + c$$
$$b3 = (a1)2 + (\lambda \ + 1) \ a3$$

Point compression permits the points on an elliptic curve additionally to be denoted with small amounts of data. In implementations of smart cards point compression is important because it lowers down not only the space of storage for keys on card, but also the huge number of data that requires to be transformed to and from the card. It can be accommodated with disregarded computation using GF (2m), but can cause implementations of GF (m) considerably. The hardware implementations of GF (2m) provide essential area size and performance benefits over hardware implementations of GF (m). Smart cards need several varied services of cryptographic with vastly fast performance may need coprocessors of cryptography. The coprocessor is configured to effective as possible GF (2m) may gain less space on the cost and smart card and may offer superior performance to an implementation of GF (m) [18].

6. CONCLUSION

For smart cards elliptic curve cryptography is the most comfortable cryptosystem. Implementing smart cards using elliptic curve cryptography saves cost; time and area. Especially smart cards that are implemented over Galois field

of order 2m, where m is a prime number, are very efficient in terms of performance as well as security. GF (m) and GF (2m) are the two extensive methods of implementation of smart cards currently in practice. However, with research being conducted in this area, to a great extent, there must be new methods of implementing Elliptic curve cryptography in the near future.

REFERENCES

1. Silverman J H, *The arithmetic of Elliptic curves, Springer, Germany, (2009).*
2. Pachgare V K, *Cryptography and Information Security, PHI Learning Private Limited, New Delhi, p 154, (2009).*
3. Tilborg H C V A and Jajodia S, *Encyclopedia of Cryptography and Security, Springer, New York, p 397, (2011).*
4. Stavroulakis P and Stamp M, *Handbook of Information and Communication Security, Springer, Germany, p 35, (2010).*
5. Yalcin S B O, *Radio Frequency Identification: Security and Privacy Issues, Springer, Germany, p 3-11,(2010).*
6. Zheng P and Lionel M N, *Smart phone and next generation mobile computing, Morgan Kauffmann Publishers, UK, p 354, (2006).*
7. Mayes K E and Markantonakis K, *Smart cards, tokens, security and applications, Springer, Germany, p 1-2, (2009).*
8. Cranor L F and Garfinkel S, *Security and usability: designing secure systems that people can use, O'Reily Media Inc., USA, p 229, (2005).*
9. Hansmann U, *Smart card application development using Java: with 98 figures, 16 tables and a multi function smart card, Springer, Germany, p 13-14, (2002).*
10. Chen Z, *Java Card technology for Smart Cards: architecture and programmer's guide, Sun Microsystems Inc., USA, p 3, (2000).*
11. American Heritage Dictionary, *High definition: an A to Z guide to personal technology, Houghton Mifflin, USA, p 292, (2006).*
12. Beiske, Lee, Yim and Yu, *E Octopus in Hong Kong – A Feasibility Study, GRIN Verlag, Germany, p 9, (2005).*
13. Ranki W and Effing W, *Smart Card Handbook, John Wiley & Sons, UK, p 20, (2010).*
14. Atkins D, *The Smart Card Report, Elsevier, UK, p 271, (2003).*
15. Hendry M, *Smart card security and applications, Artech house, USA, p 88-89, (2001).*
16. Jurgensen T M and Guthery S B, *Smart cards: the developer's toolkit, Pearson Education, New Jersey, p 34, (2002).*
17. Tipton H F and Krause M, *Information Security Management Handbook, CRC Press, USA, p 1064-1065, (2007).*
18. Stajano F, *Security and privacy in ad-hoc and sensor networks, Springer, (2007).*

Architectural Strategies in Cold Regions to Create Sustainability in Residential Spaces

Mansour Nikpour[1], Farhad Kazemian[2], Nasim Bahmani[3]

[1, 2, 3](Architecture Department, Islamic Azad University, Bam Branch, Iran)

ABSTRACT

Disregarding the excessive consumption of energy to provide comfort in the residential units, especially in cold and mountainous regions, is one of the problems of residential complexes in Iran. Although a lot of researches have been done in this field, the inattention to this matter not only causes energy crisis in the future but makes the residential complexes as the biggest environment polluter sources. On the other hand, the traditional architects of Iran have provided the residential areas, especially by considering the hard conditions of cold regions, with comfort by applying simple and available techniques and by the minimum use of fossil fuels.

Therefore at the first of this research there's a review on the regional specifics and climatic properties of cold and mountainous regions and then through the observation and former researches the specifics and strategies of residential areas architecture are discussed and reviewed from the viewpoint of providing comfort conditions by minimum amount of fossil energy, so that the effect of each feature in reducing the energy consumption could be evident. As a result, the possibility of creating sustainable residential areas are provided in the future by applying the traditional architecture strategies

Keywords - Sustainable Architecture, Residential Areas, Comfort Conditions, Architectural Techniques

I. INTRODUCTION

Mountains and High Plateau Region is one of the four climatic divisions of Iran [1].

Alborz and Zagros mountain chains separate the central areas of Iran from the Caspian Sea in the north and Mesopotamia in the west. There are also single mountains in the center and east of Iran including Taftan Mountain, Shirkooh. The western highlands, that surround the western highlands, that surround the western slopes of Iranian central plateau mountain chains and all the Zagros mountains, are considered as cold regions of the country. The climatic generalities of this region are as follows:
Severe chill in winter and temperate weather in summer;

Extreme difference between day and night temperatures;
Heavy Snows;
Low humidity of weather.
The average weather temperature in the warmest month of year is more than 10 degree Centigrade and less than 3 degree Centigrade below zero in the coldest month of year. The temperature fluctuation during day and night is more in the mountainous regions. The valleys in this region are very warm in summer and temperate in winter. Rate of sunshine in this region is high in summer and very low in winter. The winters are long, cold and severe, and the earth is covered by ice for several months of year; the spring that separates the winter and summer stays for a short time. The cold weather starts from the first of October and continues nearly to the end of April. All over this region, from Azarbayejan to Fars Province, is severely cold in winters. In these regions the amount of precipitation is low in summer and high in winter and it's mostly snowfall. The continuous snowfalls cover most of tops. There's always snow in the heights higher than 3000 m and these highlands are the source of rivers and aqueducts of the country.

Snowfall in the north and northwestern areas of this region is more than western south areas. Despite the great deal of precipitations, the humidity of this region is low. The western mountain ranges also prevent the penetration of humid Mediterranean weather to Iranian plateau and keep the moisture in their own slopes. Despite the heavy weather of the northern regions of Iran and the coasts of the Caspian Sea, which is because of the low altitude and much precipitation, the weather in the cold region is lighter and this fact decreases the use of natural air conditioning [2].

II. METHODOLOGY

This research engaged in reviewing any of the traditional residential areas of Iran via observation and evaluation of former studies and demonstrated the role of any obvious effective factor in establishing the comfort conditions with minimum energy consumption including residential areas positioning in relation with each other (city texture), orientation, general form and pattern of buildings, elements, materials, openings, dimensions and proportions; so as to create sustainable residential areas by help of traditional architects' strategies in design and construction.

III. FINDING AND DISCUSSION

3.1 CITY CONTEXT CHARACTERISTICS

The city texture of the cold and mountainous region has been developed in order to cope with extreme cold. Specifics of urban and rural texture in this region are as follows:

1. Compact and intensive texture

2. Small and enclosed areas

3. Taking advantage of the sun and earth directions (As determining factors for establishment and expansion of the city and its appearance)

4. Narrow passages along the ground level[1].

Considering the cold and mountainous region in Iran, the buildings are constructed compact and joined together to decrease the contact surface between warm residential areas and outer cold environment, so that the heat dissipation and draught is prevented. The building are also located side by side in a way that enclose each other and the urban spaces become as small as possible to reduce the cold wind penetration into urban spaces and let the heat reflection from the outer surface of warm walls of the buildings decreases and moderates the coldness of weather in small and enclosed urban areas "Fig." 1.

"Fig. "1.   View from Mountain toward residential context

The other obvious point in this type of cities is design of narrow passages to take more advantage from heat and prevent cold and heat exchange. In this type of region the residential areas are usually established in the middle of heights slope toward the south and in or on the ground. This is done for the following reasons: first, to enhance the heat capacity of northern walls and increase the interior volume in relation to exterior surface. Secondly, penetration of the heavy cold weather to the valleys increases the chill severity during night. Thirdly, the northern front of the mountain is usually under shadow and is cold, while the cities should be built toward the valley and in the sun to make the maximum use of sunshine. Fourthly, because of the excess roughness and severe winds on the height of mountain, and on the other hand, the availability of water resources and running rivers in the bottom of heights, it's not correct to establish the urban texture on the height of mountain "Fig." 2 [2, 3, 4].

Also for preventing cold wind, cities are located in the middle of hillside In this area the central yard is

smaller than central yard of buildings which are located in central plateau. In this region wood is used for roofs and stone is used for walls because these materials are available in these regions and have low thermal conductivity [2].
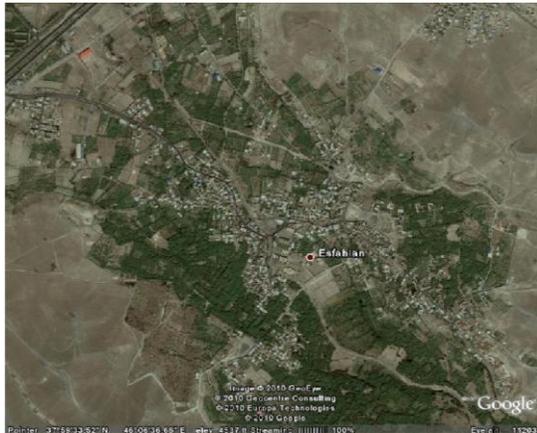
"Fig."2. Texture of Esfahlan(a city in the high plateau region)

3.2 BUILDING CHARACTERISTICS

The extreme cold weather during the greater part of year in cold and mountainous regions makes necessary for residential areas to take advantage from maximum sunshine, daily temperature fluctuation, heat protection and prevention from winter cold wind. So the form of building is designed and constructed to cope with extreme chill [3].

3.2.1 BUILDING'S FORM

The traditional houses in cold region such as central regions of Iranian plateau have a central yard and the other parts are set around this yard. The rooms located in the northern direction of yard are larger than other parts and hall or the main sitting room is also located in this direction of yard to make use of heat and direct sunshine in the cold season of winter. The southern part of building is not used much because of the short and temperate summer. So the southern rooms, eastern and western rooms, if available, are used as store room or Service areas like servants' room or bathrooms. Unlike the temperate and humid regions of southern coasts of the Capital Sea, the houses of these regions often have basement with short cellar beneath the winter room that is used

because of its cool for residence and comfort of home habitants in summer [5, 6, and 7].

3.2.2 BUILDING'S SPACES AND ELEMENTS

Since most days of year are cold or extremely cold in mountainous regions, most of daily activities are done inside the rooms. So the size of yards in these regions is slightly smaller than those of central Iranian plateau regions. Buildings in these regions have verandas but their depth is far lesser than the verandas in southern regions of country, and they don't have sitting room usage as those of the Caspian region and they're only used for snow and rain protection of building entrance. The other point is that the floor of building yards in cold region is 1 to 1.5 m below the sidewalks level to direct the current water of creeks or brook toward yard garden or water reservoir in the cellar, and on the other hand, the ground as a heat insulation surrounds the building, prevents thermal exchange between the building and environment and preserved the heat inside the building [2, 7].

3.2.3 BUILDING'S PLAN AND DIRECTION

The buildings in the cold and mountainous region have a compact plan and texture. The building formation should be in a way that reduces the contact surface with outer chill so that less heat may transfer from inside to outside. So the shapes such as cube or cubic rectangle are used to reduce the outer surface of building in relation to its inner volume and keep it in possible minimum. The buildings are established between 20 degrees to the west and 45 degrees to the east, in the wind shade of each other, out of sun shade of each other and along the north-south axis [2, 8].

3.2.4 BUILDING'S ROOM CHARACTERISTICS

In the cold and snowy regions construction of large rooms and areas inside the building should be avoided because this increases the contact surface with cold outer area and it's hard to warm this large room. So the ceiling of rooms in these regions are considered lower than the similar rooms in the other regions to decrease the room volume and the outer surface gets minimum in relation to the building volume. The low height of ceiling in halls, important

rooms, arch roofed passages and bazaar chambers are famous in these regions [2, 7 and 8].

3.2.5 BUILDING'S OPENING

Small openings in low numbers are used to prevent the thermal exchange between outside and inside of the building in these regions. If the windows are large, it's necessary to apply a shade. The openings used in the south side are larger and longer to take maximum advantage of sunshine. Also it should be prevented from settling the openings in the direction of cold winds. Double walled windows are also proper to minimize the thermal exchange. Meanwhile, the rate of internal air exchange and natural ventilation should be minimized as much as possible to prevent from breeze in the building and inner heat exit to out. In comparison with warm and dry regions, the dimensions of openings in this region are increased to make use of heat energy of sunshine [2, 5].

3.2.6 BUILDING'S WALLS

High thickness of walls, in turn, also prevents heat exchange between inner area of building and outer environment. The standards of architecture in cold and mountainous regions are nearly similar to those of warm and dry regions; and the only difference is in heat producing sources. In warm and dry regions this source is from out of building but in cold region is from inside. In these regions the wall thickness should be increased by constructional materials so that this wall may act as heat saving resource for internal area of the building. The thick walls save the heat of daily sunshine during the night and help in warming the inner space.

In local architecture of these areas, it's tried to warm the building by natural methods or by use of heater and the warmth resulted from house habitats and cookery [2, 4].

3.2.7 BUILDING'S ROOF

Traditional buildings in the northern mountain side of Alborz mountain chain have steep roofs but ones in the mountainous regions often have flat roofs. The steep roofs, if covered well, are much better than flat roofs because the rainwater is easily directed away from the roof. But if it is covered by cob (clay and straw), its resistance toward moisture, rain and specially snow is far decreased. Therefore, when there's a precipitation, the snow is shoveled from the roof at once and the roof should be rolled by a small stony roller so that the cob cover gets compact again and the holes made by water infiltration get blocked. Applying flat roofs in cold region makes no problem because keeping snow on the roof can be used as a thermal insulation against the severe chill of outer environment that is several degrees lower than snow temperature. Also the space beneath the truss, which is used as store room, is a proper insulation between inner and outer area of the building. So a double walled roof for the buildings in this region is of high importance in keeping the building heat [2].

3.2.8 BUILDING'S MATERIALS

The materials used in traditional buildings in cold and mountainous regions, like in other regions, are from the available materials there. These materials should have a good thermal capacity and resistance to keep the building warmth in its inner area. So the body of these buildings is from stone (or wood, cob mortar, adobe and bricks) and the roofing is from timber and cob. Stone and heavy resistant materials are used for building foundation, and in some parts the heavy materials are used for base course to prevent moisture. However, the buildings of these regions are generally built on the ground [5, 9, 10].

3.3 DESIGN STRATEGIES

The principles that have been thought for adaptation to climatic conditions of these regions are very important; and they are as follows:

Using common walls as much as possible and creating a heaped and compacted texture in complexes;

Preparing compressed and compact planes;

Forming the building to create shade in summer and receive proper heat in winter;

Placing heat generating spaces like kitchen in the center of building plane;

Considering non important spaces like store as heat insulator in sides or cold parts of building;

Using appropriate heat insulator in outer walls and especially in roof;

Using stony bed below the rooms to reserve extra heat in floor and release this saved heat in night or cold times[2,10].

IV. CONCLUSION

The traditional architects of Iran provide residential areas with comfort conditions by applying simple techniques. Using compact plans, minimizing the outer surface against the covered volume, applying materials with good thermal insulation and capacity and minimizing the inner air exchange and natural ventilation are among the techniques that were used by traditional Iranian architects to create comfort conditions and this case was performed with minimum use of fossil energy. Applying nature - returnable materials and minimizing the use of fossil fuels are among the aim points of sustainable architecture. Therefore, the strategies applied in Iranian traditional architecture have a great effect in creating a sustainable architecture. These techniques can also be used in the design of residential complexes in cold regions to minimize the energy consumption for providing comfort conditions. In this case, by taking advantage of these techniques the aim of sustainable architecture is achieved.

REFERENCES

[1] D. Watson, Climatic Design, Tehran University Publications, Tehran, Iran: 1994.

[2] V. Ghobadian, Climatic analysis of the tradition Iranian buildings, Tehran, Iran: Tehran University Publications, 2003.

[3] M. Kasmaei, Climate and architecture, Tehran, Iran: 2003.

[4] M.K. Pirnia, Styles of Iranian Architecture, Tehran, Iran: rush Danesh, 2004.

[5] M. K. Pirnia, Iranian Islamic architecture, Tehran, Iran: Iran University of Science and Technology, 2004

[6] M. Nikpour et al., Investigating Sustainability In Hot And Dry Climate Of Iranian Cities, Through Central Courtyard Houses , The 5th International Conference of the International Forum on Urbanism, Singapore: NUS, 2011.

[7] Gh. H. Memarian, Introduction to house typology in Iran, introvert house, Tehran, Iran: University of Science and Technology, 1995.

[8] A. Ghahramanpouri , S. Sedaghat Nia, The Effects of Climatic Factors on Urban Context and Vernacular Architecture of Cold and Temperate Regions of Iran, 3rd International Graduate Conference on Engineering, Science and Humanities, Johor, Malaysia: UTM, 2010.

[9] M. Nikpour et al., Study Of Different Climate On Urban Texture To Create Sustainable Urbanization, 3rd International Graduate Conference on Engineering, Science and Humanities, Johor, Malaysia: UTM, 2010.

[10] M. Nikpour, B. Shamsaddin, Study of Sustainable Architecture in Traditional Architecture Design Methods, 3rd International Graduate Conference on Engineering, Science and Humanities, Johor, Malaysia: UTM, 2010

Sustainable Development Using Supplementary Cementitious Materials and Recycled Aggregate

B Sarath Chandra Kumar[1], Vamsi Krishna Varanasi[1], Dr. P Saha[2]

*(Students, Department of Civil engineering, K L University, Vaddeswaram, A.P.-522502, India)
** (Associate Professor, Department of Civil engineering, K L University, Vaddeswaram, A.P.-522502, India)

ABSTRACT

Concrete is the most widely used construction material worldwide. Its popularity gives the well-known advantages, such as low cost, general availability, and wide applicability. As well as the applications of concrete in the realms of infrastructure, habitation, and transportation have greatly prompted the development of civilization, economic progress, stability and of the quality of life.

But this popularity of concrete also carries with it a great environmental cost. The billions and billions tons of natural materials are mined and processed each year, by their large volume, are bound to leave a substantial mark on the environment. Most damaging are the huge amounts of energy required to produce Portland cement as well as the large quantities of CO_2 released into the atmosphere in the process.

This paper summarizes the various efforts underway to improve the environmental friendliness of concrete to make it suitable as a "Green Building" material. For most and most successful in this regard is the use suitable substitutes for Portland cement, especially those that are by-products of industrial processes, like fly ash, ground granulated blast furnace slag, cement kiln dust, rice husk ash and wood ash.

The paper discusses some of the economic drivers which determine the degree of commercial success. Simply deposing of waste materials in concrete products is unlikely to succeed except in unusual situations. The emergence of the Green Building movement in india is already changing the economic landscape and the factors that influence resource utilization.

Keywords: Sustainable Development, Green Buildings, Supplementary Cementitious Materials, Recycled Concrete Aggregate.

I. INTRODUCTION

Presently, annual worldwide concrete production is about 12 billion tons, consuming approximately 1.6 billion tons of Portland cement, 10 billion tons of sand and rock, and 1 billion tons of water. The production of one ton of Portland cement generates approximately one ton of carbon dioxide and requires up to 7000 MJ of electrical power and fuel energy. As

Large quantities of waste materials and by-products are generated from manufacturing processes, service industries and municipal solid wastes, etc. As a result, solid waste management has become one of the major environmental concerns in the world. With the increasing awareness about the environment, scarcity of land-fill space and due to its ever increasing cost, waste materials and by-products utilization has become an attractive alternative to disposal. High consumption of natural sources, high amount production of industrial wastes and environmental pollution require obtaining new solutions for a sustainable development. During recent years there has been a growing emphasis on the utilization of waste materials and by-products in construction materials. Utilization of waste materials and by-products is a partial solution to environmental and ecological problems. Use of these materials not only helps in getting them utilized in cement, concrete, and other construction materials, it helps in reducing the cost of cement and concrete manufacturing, but also has numerous indirect benefits such as reduction in landfill cost, saving in energy, and protecting the environment form possible pollution effects. Further, their utilization may improve the microstructure, mechanical and durability properties of mortar and concrete, which are difficult to achieve by the use of only ordinary Portland cement. [1]

II. Sustainable Development

Sustainable building materials are more than recycled or reused materials and components. In order to be truly sustainable, the material must be examined from the time it is harvested as a raw material to the time it will need to be disposed of. When carrying out a Sustainable Development price is of course one consideration but more important is the environmental and health impact of the materials. Energy consumption, waste, emissions, and the resources' ability to renew itself are the most important aspects. The energy consumption of a material during harvesting, transporting, processing, and use are all considered to decide if a material is truly sustainable and therefore suitable in the construction of a Sustainable Development.

These and various related concerns led to the concept of sustainable development, which can be summarized as follows:

1. Remedy the mistakes of the past by cleaning up our contaminated water and soil.
2. Avoid the pollution of our air, water and soil, including the release of greenhouse gases into the atmosphere that are known to contribute to global warming.
3. Utilize natural resources, whether material or energy, at a rate no greater than at which they can be regenerated.
4. Find a proper balance between economic development and preservation of our environment, i.e. improve the living standard and quality of life without adversely affecting our environment. [2]

These goals describe an ideal state and are obviously difficult to achieve. Yet, we do not have much of a choice, lest the live ability of our planet take a rapid turn for the worse. As the World Earth Summits in Rio de Janeiro (1990) and Kyoto (1997) demonstrated very clearly, this worldwide problem can be solved only through concerted international efforts. The industrialized countries are called upon to reduce the emission of greenhouse gases and the wasteful use of natural resources, and the developing countries need to avoid the mistakes made by the industrialized world in the past and develop their economies using technologies that make optimal use of energy and natural materials, without polluting the environment.

It is the purpose of this article to discuss various aspects of the concrete industry, because it has a much larger impact on sustainability than many of us may realize. Concrete is by far the most widely used construction material worldwide. In fact, it is more widely used than any other material, except water. Its huge popularity is the result of a number of well-known advantages, such as low cost, general availability, and adaptability to a wide spectrum of performance requirements. But this popularity of concrete also carries with it a great cost in terms of impact on the environment. [3, 4]

1. Worldwide, over ten billion tons of concrete are being produced each year. In the United States, the annual production of over 500 million tons implies about two tons for each man, woman and child. Such volumes require vast amounts of natural resources for aggregate and cement production.
2. In addition, it has been estimated that the production of one ton of Portland cement causes the release of one ton of CO2 into the atmosphere. CO2 is known to be a greenhouse gas that contributes to global warming, and the cement industry alone generates about 7% of it.
3. The production of Portland cement is also very energy-intensive. Although the North American plants have improved their energy-efficiency considerably in recent decades to the point where this is now comparable to that of plants in Japan and Germany, it is technically next to impossible to increase that energy-efficiency much further below the current requirement of about 4 GJ per ton.
4. The demolition and disposal of concrete structures, pavements, etc., constitutes another environmental burden. Construction debris contributes a large fraction of our solid waste disposal problem, and concrete constitutes the largest single component.
5. Finally, the water requirements are enormous and particularly burdensome in those regions of the earth that are not blessed with an abundance of fresh water. The concrete industry uses over 1 trillion gallons of water each year worldwide, and this does not even include wash water and curing water.

These points and these numbers seem to indicate that the concrete industry has become a victim of its own success and therefore is now faced with tremendous challenges. But the situation is not as bad as it might seem, because concrete is inherently an environmentally friendly material, as can be demonstrated readily with a life-cycle analysis [5]. The challenges therefore reduce primarily to reducing Portland cement's impact on the environment. In other words, we should use as much concrete, but with as little Portland cement as possible. [6]

III. Challenges for concrete and cement industry for sustainable development

There are a number of ways how the concrete industry can increase its compliance with the demands of sustainable development:

1. Increased use of supplementary cementitious material. Since the production of Portland cement is energy intensive and responsible for much of the CO_2 generation, the substitution of other materials, especially those that are byproducts of industrial processes, such as fly ash and slag, is bound to have a major positive impact.
2. Increased reliance on recycled materials. Since aggregate constitutes the bulk of concrete, an effective recycling strategy will lessen the demand for virgin materials.
3. Improved durability. By doubling the service life of our structures, we can cut in half the amount of material needed for their replacement.
4. Improved mechanical properties. An increase in mechanical strength and similar properties leads to a reduction of materials needed. For example, doubling the concrete strength for strength-controlled members cuts the required amount of material in half.

5. Reuse of wash water. The recycling of wash water is readily achieved in practice and already required by law in some countries.

Implementing effective strategies to lessen the environmental impact of the concrete industry by careful use of those tools requires a concerted effort of the industry, starting with well-focused research and development. Even more important for success are economic incentives to convince industry leaders that increased incorporation of sustainable development principles is possible without adversely impacting the industry's profitability. On a less benign parallel track, political developments are underway or imminent which are likely to force the industry to change or lose market share. Bold initiatives are required that are not without risk, yet strict adherence to principles such as "we have always done it this way" is certainly counterproductive, because the world around us will change anyway.

A considerable body of literature exists on methods to improve the mechanical properties and durability of concrete. The emphasis here will be on how to make concrete a "green building material" by use of cement substitutes and recycled materials. [6]

IV. Use of Cement Substitutes for sustainable concrete

Cement is the backbone for global infrastructural development. It was estimated that global production of cement is about 1.3 billion tons in 1996. Production of every tone of cement emits carbon dioxide to the tune of about 0.87 ton. Expressing it in another way, it can be said that 7% of the world's carbon dioxide emission is attributable to Portland cement industry. As we all know that carbon dioxide is one of the significant green house gas and it contribution to the environmental pollution is very high. The ordinary Portland cement also consume natural resources like limestone etc., that is why we cannot go on producing more and more cement and there is a need to economize the use of cement. One of the practical solutions to economize cement is to replace cement with supplementary cementitious materials. [7]

Cement is the key component of concrete that binds the other components together and gives the composite its strength. A considerable amount of work has been reported in the literature on how to use waste products of combustion or industrial processes as supplementary cementitious materials [4, 8, 9]. Because of their cementitious or pozzolanic properties these can serve as partial cement replacement. Ideally, the development of such materials serves three separate purposes simultaneously. On the one hand, waste byproducts have an inherent negative value, as they require disposal, typically in landfills, subject to tipping fees that can be substantial. When used in concrete, the material's value increases considerably. The increase in value is referred to as "beneficiation".

As this supplementary cementitious material (SCM) replaces a certain fraction of the cement, its market value may approach that of cement. A second benefit is the reduction of environmental costs of cement production in terms of energy use, depletion of natural resources, and air pollution. Also, the tangible as well as intangible costs associated with landfilling the original waste materials are eliminated.

Finally, such materials may offer intriguing additional benefits. Most concrete mixes can be engineered such that the SCM will give the mix certain properties (mechanical strength, workability, or durability) which it would not have without it. It is the challenge for the concrete technologist when developing a mix design, to combine these three different goals in an optimal way such that the economic benefits become transparent. The key task is to turn waste material with a large inherent negative value into a potentially valuable product. The increase in value should be both real, in terms of converting a liability into a commodity with an increased market value, as well as intangible in terms of reduced environmental costs. The fundamental challenge for the researcher is to identify waste materials with inherent properties that lend themselves to such beneficiation. Below, a few examples shall be mentioned.

A primary goal is a reduction in the use of Portland cement, which is easily achieved by partially replacing it with various cementitious materials, preferably those that are byproducts of industrial processes. The best known of such materials is **fly ash**. Fly ash is finely divided residue resulting from the combustion of powdered coal and transported by the flue gases and collected by electrostatic precipitator. In U.K. it is referred as pulverised fuel ash. Fly ash is the most widely used pozzolanic material all over the world. Fly ash was first used in large scale in the construction of Hungry Horse dam in America in the approximate amount of 30 per cent by weight of cement. Later on it was used in Canyon and Ferry dams etc. In India, Fly ash was used in Rihand dam construction replacing cement up to about 15 percent [7]. As shown in Table 1 [6], the utilization rates vary greatly from country to country, from as low as 3.5% for India to as high as 93.7% for Hong Kong. The relatively low rate of 13.5% in the US is an indication that there is a lot of room for improvement.

Table 1: Coal-Ash Production and Utilization [6]

Country	Million Tons Produced	Million Tons Utilized	%
China	91.1	13.8	15.1
Denmark	1.3	0.4	30.8
Hong Kong	0.63	0.59	93.7
India	57.0	2.0	3.5
Japan	4.7	2.8	59.6

Russia	62.0	4.3	6.9
USA	60.0	8.1	13.5

The use of fly ash has a number of advantages. It is theoretically possible to replace 100% of Portland cement by fly ash, but replacement levels above 80% generally require a chemical activator. We have found that the optimum replacement level is around 30%. Moreover, fly ash can improve certain properties of concrete, such as durability. Because it generates less heat of hydration, it is particularly well suited for mass concrete applications. Fly ash is also widely available, namely wherever coal is being burned. Another advantage is the fact that fly ash is still less expensive than Portland cement. May be most important, as a byproduct of coal combustion fly ash would be a waste product to be disposed of at great cost

In the recent time, the importance and use of fly ash in concrete has grown so much that it has almost become a common ingredient in concrete, particularly for making high strength and high performance concrete. The new Indian Standard on concrete mix proportions (IS 10262-2009) are already incorporated fly ash as a supplementary material to cement. [10]

Extensive research had been done all over the world on the advantage of fly ash as a supplementary cementitious material. High volume fly ash concrete is a subject of current interest across the globe. ASTM broadly classifies fly ash into two classes. Class F and class C. Class F Fly ash normally produced by burning anthracite or bituminous coal and has pozzolanic properties only. Class C Fly ash normally produced by burning lignite or sub-bituminous coal and can possesses pozzolanic as well as cementitious properties. [7]

The employment of fly ash in cement and concrete has gained considerable importance because of the requirements of environmental safety and more durable construction in the future. The use of fly ash as partial replacement of cement in mortar and concrete has been extensively investigated in recent years.it is known that the fly ash is an effective pozzolan which can contribute the properties of concrete. Fly ash blended concrete can improve the workability of concrete compared to OPC. It can also increase the initial and final setting time of cement pastes. Fly ash replacement of cement is effective for improving the resistance of concrete to sulfate attack expansion. The higher is the compressive strength of concrete,the lower is the ratio of splitting tensile strength to compressive strength. Finally, it is knowing that the properties of concrete are enhanced when the substitution of Portland cement was done by fly ash. [10]

Ground granulated blast furnace slag (GGBS) is a by-product of the manufacturing of iron in a blast furnace where iron ore, limestone and coke are heated up to 1500◦C.When these materials melt in the blast furnace, two products are produced i.e molten iron, and molten slag. The molten slag is lighter and floats on the top of the molten iron. The molten slag comprises mostly silicates and alumina from the original iron ore, combined with some oxides from the limestone. The process of granulating the slag involves cooling the molten slag through high-pressure water jets. This rapidly quenches the slag and forms granular particles generally not larger than 5mm in diameter. The rapid cooling prevents the formation of larger crystals, and the resulting granular material comprises some 95% non-crystalline calcium-aluminosilicates. The granulated slag is further processed by drying and then ground to a very fine powder, which is GGBS (ground granulated blast furnace slag) cement It is another excellent cementitious material.[14]

Wainwright and Ait-Aider (1995) examined the influence of the composition of OPC and the addition of up to 70% GGBS on the bleed characteristics of concrete and conclude that the partial replacement of OPC with 40% and 70% of GGBS. GGBS led to increases in the bleeding of the concretes, like fly ash, also GGBS can improve many mechanical and durability properties of concrete and it generates less heat of hydration.[11]

Babu and Kumar (2000) determined the cementitious efficiency of GGBS in concrete at various replacement percentages (10–80%) through the efficiency concept by establishing the variation of strength to water-to-cementitious materials ratio relations of the GGBS concretes from the normal concretes at the age of 28 days. The 28-day compressive strength of concretes containing GGBS up to 30% replacement were all slightly above that of normal concretes, and at all other percentages, the relationships were below that of normal concretes. It was also observed that the variations due to the different percentages of slag replacement were smaller than the corresponding variations in the case of fly ash. The result showed that the slag concretes based on overall efficiency factor (k), will need an increase of 8.6% for 50% replacement and 19.5% for 65% replacement in the total cementitious materials for achieving strength equivalent to that of normal concrete at 28 days.[12]

Rice-husk is an agricultural by-product material. When rice-husk is burnt rice-husk ash (RHA) is generated. RHA is highly pozzolanic material. The non-crystalline silica and high specific surface area of the RHA are responsible for its high pozzolanic reactivity. RHA has been used in lime pozzolana mixes and could be a suitable partly replacement for Portland cement.[13]

RHA concrete is like fly ash/slag concrete with regard to its strength development but with a higher pozzolanic activity it helps the pozzolanic reactions occur at early ages rather than later as is the case with other replacement cementing materials. [14]

The employment of RHA in cement and concrete has gained considerable importance because of the

requirements of environmental safety and more durable construction in the future. The use of RHA as partial replacement of cement in mortar and concrete has been extensively investigated in recent years.

RHA blended concrete can decrease the temperature effect that occurs during the cement hydration. RHA blended concrete can improve the workability of concrete compared to OPC. It can also increase the initial and also final setting time of cement pastes. Additionally, RHA blended concrete can decrease the total porosity of concrete and modifies the pore structure of the cement, mortar, and concrete, and significantly reduce the permeability which allows the influence of harmful ions leading to the deterioration of the concrete matrix. RHA blended concrete can improve the compressive strength as well as the tensile and flexural strength of concrete. RHA helps in enhancing the early age mechanical properties as well as long-term strength properties of cement concrete.

Partial replacement of cement with RHA reduces the water penetration into concrete by capillary action. RHA replacement of cement is effective for improving the resistance of concrete to sulfate attack. The sulfate resistance of RHA concrete increases with increasing the RHA replacement level up to 40%. Substitution of RHA has shown to increase the chemical resistance of such mortars over those made with plain Portland cement. Incorporation of RHA as a partial cement replacement between 12% and 15% may be sufficient to control deleterious expansion due to alkali-silica reaction in concrete, depending on the nature of the aggregate. It can be known that the use of rice husk ash leads to enhanced resistance to segregation of fresh concrete compared to a control mixture with Portland cement alone. Also RHA can significantly reduce the mortar-bar expansion.Finally showed that the mechanical properties of concrete are enhanced when the substitution of Portland cement was done by RHA. [13]

Wood ash is the residue generated due to combustion of wood and wood products. It is the inorganic and organic residue remaining after the combustion of wood or unbleached wood fiber. [15]

Abdullahi (2006) reported the compressive strength test results of wood ash (WA) concrete. He used wood ash as partial replacement of cement in varying percentages (0, 10, 20, 30, and 40%) in concrete mixture proportion of 1:2:4. Tests were conducted at the age of 28 and 60 days.The results showed that the specimens containing 0% wood ash had the highest compressive strength. The mixture containing 20% wood ash had higher strength than that containing 10% wood ash at 28 and 60 days. This was due to the fact that the silica provided by 10% wood ash was inadequate to react with the calcium hydroxide produced by the hydration of cement. Increase in wood ash content beyond 20% resulted in a reduction in strength at 28 and 60 days. [16]

Naik et al. (2002) investigated the compressive strength of concrete mixtures made with wood ash up to the age of 365 days. Wood ash content was 5, 8, and 12% of the total cementitious materials. Figure 9.1 shows the compressive strength results. Based on the results, they concluded that: (i) control mixture (without wood fly ash) achieved strength of 34 MPa at 28 days and 44 MPa at 365 days; (ii) strength of concrete mixtures containing wood fly ash ranged from 33 MPa at 28 days and between 42 and 46 MPa at 365 days; and (iii) inclusion of wood fly ash contributed to the strength development of concrete mixtures, even as the cement content was decreased by about 15%. This indicates contribution of wood fly ash to pozzolanic activity. [17]

Udoeyo et al. (2006) determined the compressive strength of concrete made with varying percentages (5, 10, 15, 20, 25, and 30 by weight of cement) of waste wood ash (WWA). They reported that compressive strength generally increased with age but decreased with the increase in the WWA content. Comparisons of the strength of WWA concrete with those of the control (plain) concrete of corresponding ages showed that the strength of WWA concrete was generally less than that of the plain concrete. A possible explanation for this trend is that the WWA acts more like filler in the matrix than as a binder. Thus, increasing the ash content led to an increase in the surface area of the concrete filler to be bonded by the same amount of cement as that of the control. [18]

Cement kiln dust (CKD) is a by-product of cement manufacturing. It is a fine powdery material similar in appearance to Portland cement. EL-Sayed et al. (1991) investigated the effect of cement kiln dust (CKD) on the compressive strength of cement paste and corrosion behavior of embedded reinforcement. They reported that up to 5% substitution of CKD, by weight of cement had no adverse effect on the strength of the cement paste and on reinforcement passivity. Batis et al. (2002) also reported the increase in compressive strength and corrosion resistance of the mix when CKD and blast furnace slag were added in proper ratio in ordinary Portland cement. [15, 19,20]

Shoaib et al. (2000) studied the effect partial substitution of ordinary Portland cement (OPC) and blast furnace slag cement with cement kiln dust (CKD) on the compressive strength of concrete and also determined the optimum quantity of CKD which could be recycled in the manufacture of these types of cements.

Shoaib et al. (2000) studied the effect partial substitution of ordinary Portland cement (OPC) and blast furnace slag cement with cement kiln dust (CKD) on the compressive strength of concrete and also determined the optimum quantity of CKD which could be recycled in the manufacture of these types of cements. The CKD contained the mixture of raw feed, partially calcined cement clinker, and condensed volatile salts.Percentages of replacement of CKD as

ratio to cement used were 0, 10, 20, 30, and 40%. The mixes have same mix proportion (1 Cement: 1.9 Sand: 3.52 Gravel and 0.5 W/C ratio), and the cement content used in the mixes was 350Kg/m3. Compressive strength was determined at the age of 1, 3, and 6 months. Based on the test results, they reported that control mix (0% CKD) achieved compressive strength of 27, 28.5 and 32.0MPa at the age of 1, 3, and 6 months, respectively. Compressive strengths of concrete mixes decreased with the increase in CKD percentage at all ages. [21]

V. RECYCLED AGGREGATE

Aggregate constitutes approximately 70% of concrete volume. Worldwide, this amounts to billions of tons of crushed stone, gravel, and sand that need to be mined, processed, and transported every day. [2]
The substitute material that comes to mind first is recycled concrete. Construction debris and demolition waste constitute 23% to 33% of municipal solid waste, and demolished concrete contributes the largest share of this waste material (Recycling Concrete Saves Resources, Eliminates Dumping). [22]

Concrete debris is probably the most important candidate for reuse as aggregate in new concrete. On the one hand, vast amounts of material are needed for aggregate. On the other hand, construction debris often constitutes the largest single component of solid waste, and probably the largest fraction of this is concrete. Using such debris to produce new concrete conserves natural resources and reduces valuable landfill capacity at the same time. In Europe and Japan, such recycling is already widely practiced [23, 24], whereas in the US, it is being accepted only slowly, because the economic drivers are not yet strong enough. But they are improving. The disposal of demolished concrete involves costs, which are likely to go up. Available sources of suitable virgin aggregate are being depleted, such as gravel pits on Long Island, and opening new sources of virgin material is getting increasingly difficult because of environmental concerns. Since the cost of transportation is the main component of the cost of bulk material like sand and gravel, it may not take much of a shift to turn the economics in favor of recycling and reuse.
Turning recycled concrete into useful or even high-quality aggregate poses well-known technical challenges [23]. There are contaminants to be dealt with, high porosity, grading requirements, as well as the large fluctuations in quality. Not all applications require high-strength concrete, though. Recycled concrete aggregate is likely to be quite adequate for some projects, while for others, a blend of new and recycled aggregate may make most economic and technical sense. [6]

VI Changing Political Landscape in India

There are signs that the public attitude towards sustainable development is changing. "Green building design" principles are finding their way into design practice, spearheaded by the architectural community. The Indian Green Building Council has developed a rating system with the help of United States green building council as a guide for green and sustainable design. This system, called "Leadership in Energy & Environmental Design" (LEED), has become a standard adopted by several governmental agencies in its original form or some modified versions of it. It assigns points in seven different categories:

1. Sustainable Sites, 28 possible points
2. Water Efficiency, 10 possible points
3. Energy & Atmosphere, 37 possible points
4. Materials & Resources, 14 possible points
5. Indoor Environmental Quality, 15 possible points
6. Innovation & Design Process, 6 possible points
7. Regional Priority, 4 possible points

In order to become "certified", a project requires at least 40-49 points. Projects with 50-59 points are "Silver"-rated, those with 60-69 points are "Gold"-rated, and to reach the highest rating of "Platinum", 80 points and above are required.[24] Means and methods to increase the number of points for a concrete building can be found elsewhere [5]. Several industry-wide efforts are currently underway to develop guides for the industry, to not only increase the number of LEED-points, but also to improve the environmental friendliness of concrete construction across the board. Here it suffices to point out that under the current system, only a rather small number of points can be earned by making concrete more environmentally friendly. For example, in a mix design that contains 15% cementitious material, the replacement of 30% of Portland cement by fly ash will introduce only 4.5% recycled material. The reward in terms of LEED-points in no way reflects the gain in environmental friendliness, as measured by the reduction of CO_2 generation and energy consumption. [6]

VII Conclusion

Virgin materials have a quality control advantage over recycled materials. But the economic feasibility of recycling will increase in time, as virgin materials become more and more scarce and the disposal costs of construction debris and other waste materials keep increasing. The economic feasibility of recycling depends largely on the application. Concrete and cement industry can contribute to sustainable development by adopting supplementary cementitious materials, recycled aggregate to save natural resources, energy, reducing CO_2 emissions, and protect the environment and can improve its record with an increased reliance on recycled materials and in particular by replacing larges percentages of

Portland cement by byproducts of industrial processes. This will help our sustainable and green environment.

References:

[1]. Siddique Rafat and Mohammad Iqbal Khan, Supplementary Cementing Material, *Springer Heidelberg Dordrecht*, London, New York, 2011.

[2]. Meyer, Concrete and Sustainable Development, *Special Publication ACI 206*, 1-12,2002.

[3]. Mehta P.K,Greening of the Concrete Industry for Sustainable Development,*Concrete International*, 23-28,2002.

[4]. Malhotra V.M., Role of Supplementary Cementing Materials in Reducing Greenhouse Gas Emissions, *Concrete Technology for a Sustainable Development in the 21st Century*, O.E. Gjorv and K. Sakai, eds., E&FN Spon, London, 2000.

[5]. Van Geem M.G. and M.L. Marceau,Using Concrete to Maximize LEEDS Points, *Concrete International*, 2002.

[6]. Meyer,Concrete as a green building material, Invited Lecture, *Proceedings of the Third Int. Conference on Construction Materials*, ConMat'05, Vancouver, Aug. 22-25,2005.

[7]. Shetty.M.S, Concrete Techonology, *S. Chand & Company Ltd*. New Delhi,2005.

[8]. Fly Ash, Slag, Silica Fume and Other Natural Pozzolans, *Proceedings, 6th International Conference, Special Publication 178,* American Concrete Institute, Farmington Hills, MI,1998.

[9]. Fly Ash, Slag, Silica Fume and Other Natural Pozzolans, *Proceedings, 5th International Conference, Special Publication 153*, American Concrete Institute, Farmington Hills, MI,1995.

[10]. Sarath Chandra Kumar. Bendapudi, P. Saha, Contribution of Fly ash to the properties of Mortar and Concrete, *International Journal of Earth Sciences and Engineering*, ISSN 0974-5904, Volume 04, No 06 SPL, pp 1017-1023,2011.

[11]. Wainwright PJ, Ait-Aider H,The influence of cement source and slag additions on the bleeding of concrete, *Cement and Concrete Research* 25 (7): 1445–1456,1995.

[12]. Babu KG, Kumar VSR, Efficiency of GGBS in concrete, *Cement and Concrete Research,* 30: 1031–1036,2000

[13]. Alireza Naji Givi, Suraya Abdul Rashid, Farah Nora A. Aziz, and Mohamad Amran Mohd Salleh, 2010,Contribution of rice husk ash to the properties of mortar and concrete: A Review, *Journal of American Science*, 6(3), 2010.

[14]. Molhotra V.M.,Fly Ash, Slag, Silica Fume, and Rice Husk Ash in Concrete: A review, *Concrete International*, 15(4): 2-28,1993.

[15]. Siddique Rafat, Waste Materials and By-Products in Concrete, *Springer-Verlag Berlin Heidelberg*, 2008.

[16]. Abdullahi M, Characteristics of wood ash/OPC Concrete, *Leonardo Electronic Journal of Practices and Technologies,* 8: 9–16,2006.

[17]. Naik TR, Kraus RN, Siddique R, *Demonstration of manufacturing technology for concrete and CLSM utilizing wood ash from Wisconsin,* Report No. CBU-2002-30, Report for Year 1 activities submitted to the Wisconsin Department of Natural Resources, Madison, WI, for Project # 01-06 UWM Center for By-Products Utilization, Department of Civil Engineering and Mechanics, University of Wisconsin-Milwaukee, Milwaukee, pp. 124, 2002.

[18]. Udoeyo FF, Inyang H, Young DT, Oparadu, EE, Potential of wood waste ash as an additive in concrete, *Journal of Materials in Civil Engineering*,18 (4): 605–611,2006.

[19]. EL-Sayed HA, Gabr NA, Hanafi S, Mohran MA,Reutilization of by-pass kiln dust in cement manufacture, *International Conference on Blended Cement in Construction*, Sheffield, UK,1991.

[20]. Batis G, Rakanta E, Sideri E, Chaniotakis E, Papageorgiou A,Advantages of simultaneous use of cement kiln dust and blast furnace slag,*International Conference on Challenges of Concrete Construction*,University of Dundee, Dundee, UK,2002.

[21]. Shoaib MM, Balaha MM, Abdel-Rahman AG,Influence of cement kiln dust substitution on the mechanical properties of concrete, *Cement and Concrete Research* 30 (3): 371–337,2000.

[22]. Recycling Concrete Saves Resources, Eliminates Dumping, *Environmental Council of Concrete Organizations*, Skokie, IL,1997

[23]. Hansen, T.C., (Ed.), Recycling of Demolished Concrete and Masonry, *RILEM Report 6*, Chapman and Hall, London,1992.

[24]. ACI Committee 555, Removal and Reuse of Hardened Concrete, *American Concrete Institute*, Report ACI 555R-01,2001.

[25]. LEED 2011 for India- Core & Shell, *Indian Green Building Council*, November, 2011.

Quantum Inspired Evolutionary Technique for Optimization of End Milling Process

Rajat Setia*, K. Hans Raj*, Suren N. Dwivedi**

(*Mechanical Engineering Department, Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra, India)
(**Mechanical Engineering Department, University of Louisiana at Lafayette, U.S.A.)

ABSTRACT

In this paper an attempt is made to develop a new Quantum Inspired Evolutionary Technique (QIET) that is general, flexible and efficient in solving single objective constrained optimization problems. It generates initial parents using quantum seeds. It is here that QIET incorporates ideas from the principles of quantum computation and integrates them in the current frame work of Real Coded Evolutionary Algorithm (RCEA). It also incorporates Simulated Annealing (SA) in the selection process of Evolutionary Algorithm (EA) for child generation. In order to test this algorithm on domain specific manufacturing problems, Neuro-Fuzzy (NF) modeling of end milling process is attempted and the NF model is incorporated as a fitness evaluator inside the QIET to form a new variant of this technique, i.e. Quantum Inspired Neuro Fuzzy Evolutionary Technique (QINFET) and is effectively applied for process optimization of end milling process. The optimal process parameters obtained by QINFET correlates better than those reported in literature. The proposed methodology using QINFET is a step towards meeting the challenges posed in intelligent manufacturing systems and opens new avenues for parameter estimation and optimization.

Keywords – **End Milling**, NF Modeling, QIET, QINFET.

I. INTRODUCTION

The last two decades have witnessed tremendous growth in the application of stochastic search techniques. The primary reason for this is that these are well suited to the concurrent manipulation of models of varying resolution and structure. This is due to their ability to search non-linear space without gradient information or a prior knowledge relating to model characteristic. The most important stochastic search techniques that have been popular are Evolutionary Strategies (ES), Genetic Algorithms (GA), Simulated Annealing (SA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Immune Algorithm, Tabu Search (TS) and Quantum-inspired Evolutionary Algorithms (QIEA).

Many efforts have been made by researchers to overcome limitations of earlier algorithms such as slow and premature convergence by establishing a good balance between exploitation and exploration. One such effort resulted in the hybridization of Evolutionary Algorithms (EA) with other heuristics such as simulated annealing, local search, tabu search, hill climbing, dynamic programming, greedy random adaptive search procedure and quantum computing. This hybridization resulted in the improvement of performance in terms of convergence speed and quality of the solutions obtained by EA [1, 2].

Hans Raj et al. [3] have proposed a hybrid Evolutionary Computational Technique (ECT) by combining GA and SA. It is a hybrid scheme which incorporates a real–coded GA to provide multi-point search along with simulated annealing method to overcome local convergence and the problem of multiple minima. This technique provides more rapid and robust convergence on many function optimization problems. Two levels of competition are introduced between the strings in the population to ensure that only the better strings continue in the population. The concept of "Acceptance Number" is introduced to ensure that more computational effort is devoted to search in "better" regions of the search space. Constraints are handled by the use of the concept of penalty functions and by better coding.

This paper proposes a new variant of Real Coded Quantum Evolutionary Algorithm (RCQEA) using quantum computation principles to seed initial populations in the current framework of ECT as proposed by Hans Raj et al. [3], namely, QIET, which is more suitable than ECT for a wide range of real-world numerical optimization problems. To verify its effectiveness it has been applied to optimize end milling cutting conditions to obtain the best compromise between two critical machining-related values: surface roughness and machining time. Spindle speed, feed rate, radial depth of cut and tolerance were optimized, while one of the two key performance values was kept in the desired range and the other one was minimized.

The paper is organized as follows. The background material for quantum computation is described in section II. In section III QIET is described in detail. In section IV Neuro Fuzzy modeling of end milling process is detailed and in section V the new Quantum Inspired Neuro Fuzzy Evolutionary Technique is explained and applied to end milling process.

II. INTRODUCTION TO QUANTUM COMPUTATION

A classical bit can only be in one of two states, 0 or 1 but according to the principles of quantum computation a qubit (or quantum bit) may be in the '1' state, in the '0' state, or in any superposition of the two. The state of a qubit can be represented as:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where α and β are complex numbers that specify the probability amplitudes of the corresponding states. α^2 gives the probability that the qubit will be found in the '0' state and β^2 gives the probability that the qubit will be found in the '1' state. Normalization of the state to unity guarantees

$$|\alpha|^2 + |\beta|^2 = 1$$

The interesting part is that until the qubit is measured it is effective in both states. The probability of measuring the answer corresponding to an original 0 bit is α^2 and the probability of measuring the answer corresponding to an original 1 bit is β^2 [4, 5].

III. QUANTUM INSPIRED EVOLUTIONARY TECHNIQUE (QIET)

In QIET the idea is to seed the initial population with a quantum approach in the framework of ECT. The algorithm concentrated its search only in the permissible regions of the search space using penalty approach. QIET search technique starts out with a guess of N grandparents, chosen at random in the search space. Initially each grandparent generates a number of quantum parents.

Number of quantum parents is chosen to be 10. Larger population sizes might yield further improvement in the results obtained but would entail higher computational effort. All the variables are encoded as floating point numbers. The method selects grandparents as random numbers in the range (0, 1) for each element of chromosome. Considerably distant points in the solution space are generated as grandparents to avoid any domination of particular schemata during the initialization process. Initially all grandparents generate an equal number of quantum parents. The advantage of using a quantum seeded generation is that a number of quantum parents are generated with each grandparent without using GA. The idea is taken from the point of view that in a parental string any value generated will either be bigger or smaller than another randomly generated number, which is known as probability of finding this string value in a particular state. If the string value is less than the random probability it is retained as such else it is changed as:

$$\sqrt{[1-(String\,Value)^2]}$$

This idea is in accordance with the principles of quantum computation as described in section II, which states that the probability of a qubit to be in any state satisfies the condition $|\alpha|^2 + |\beta|^2 = 1$. Thus in a single pass numerous quantum parents can be generated with a single grandparent. Each quantum parent is checked over its functional value and constraints violation. From these quantum parents further parents are selected. A quantum parent is made a parent only when it clears a criterion of the sum of penalties for all the constraints violated. Thus the total number of parents selected varies for each iteration and also varies with a particular run of the program. Now these parents are sent into ECT which is combined GA/SA and is used for further child generation.

Initially all parents generate an equal number of children given by $m(i) = M$. A reasonable value of M is taken as 10. A higher value of M results in a more exhaustive search with a corresponding increase in computational effort. The total number of children in a generation is fixed and is given by:

$$TC = \sum_{i=1}^{N} m(i)$$

For each parent i, mates are selected from the other parents at random and cross-over is applied to generate $m(i)$ children.

For each family a blend cross-over operator (BLX-α) based on the theory of interval schemata is employed in the study. BLX-α operates by randomly picking a point in the range $(p_1 - \alpha (p_2 - p_1), p_2 + \alpha (p_2 - p_1))$ where p_1 and p_2 are two parent points and $p_1 < p_2$. In a number of test problems BLX-0.5 performed better than the BLX operators with any other α values and has, therefore, been used. Mutation is not employed.

The best child (with minimum objective value) out of the children generated from the same parent is found. The best child then competes with its parents to survive in the next generation. If the best child is better than its parent, it is accepted as a parent in the next generation. If the best child is worse than its parent then Boltzmann criterion is applied before the child be accepted.

As in SA, the selection of temperatures is such that initially the probability of acceptance of a bad move, i.e. when the best child is worse than the parent is high (approximately 1) but as the temperatures are successively lowered through a cooling schedule this probability is decreased until, at the end, the probability of accepting a bad move is negligible (approximately 0). Logarithmic cooling schedule is adopted

in this work. Such a strategy enables the technique to seek the global optimum without getting stuck in any local optimum. The initial and final temperatures are calculated as follows:

A bad move is accepted according to the Boltzmann Criterion. Initially the probability of accepting a bad move is approximately one i.e.
$exp\ (-\Delta X_{average}/T_1)=0.99$

and finally $exp\ (-\Delta X_{average}/T_{MAXIT})=0.0001$

Therefore, $T_1=-\Delta X_{average}/log\ (0.99)$

$T_{MAXIT} = -\Delta X_{average}/log\ (0.0001)$

where T_1 is the initial temperature, T_{MAXIT} is the final temperature, $\Delta X_{average}$ is the average difference between the objectives X for any two neighborhood points in the search space. This average is calculated over a number of chromosomes.

The number of children that are generated in the next generation is proportional to a parameter called the acceptance number. This number provides a measure of the goodness of solutions in the vicinity of the current parent. The number is computed by sampling the search space around the current parent and counting the number of good samples out of the total samples as per steps 13 to 15 of the pseudo-code. This strategy enables the algorithm to focus search on the better regions of the search space.

For highly constrained problems, infeasible solutions may occupy a relatively big portion of the population. The penalty technique is perhaps the most common technique used to handle infeasible solutions in the constrained optimization problems. In essence, this technique transforms the constrained problem into an unconstrained problem by penalizing infeasible solutions, in which a penalty term is added to the objective function for any violation of the constraints. The major concern is how to determine the penalty term so as to strike a balance between the information preservation (keeping some infeasible solutions) and the selective pressure (rejecting some infeasible solutions), and avoid both under penalty and over penalty. There are no general guidelines on designing penalty function. Constructing an efficient penalty function is quite problem dependent. The same has been incorporated in the QIET algorithm in evaluating the objective function i.e.

$$Evol(X) =f(X) + \beta \times D(X)$$

where, β is the problem dependent constant, and $D(X)$ is the difference measure for constraint violation. The result of such an approach is that the infeasible strings have much worse objective function values and are eliminated from the population whereas the strings with better objective function

values survive and contribute more to the evolution of better solutions.

The number of grandparents taken depends upon what is the criterion for the selection of parents from the quantum parents. The relaxation gives us choice to start with small number of grandparents. It is seen that more constrained the criterion for the selection of parents is, the better is the convergence. The initial generation of the quantum parent ensures that parents with a better fitness value are sent into GA to produce further children. This gives GA a better convergence towards the optimum solution. The results show that the convergence is very fast. The various features explained above have been combined together to develop an optimization algorithm and is represented succinctly in the form of pseudo-code given below:

1. Generate random initial grandparent strings.
2. Generate a random probability.
3. If any of the string values is less than the random probability it is retained as such.
4. Otherwise it is changed as $\sqrt{[1-(String\,Value)^2]}$.
5. Initialize T_1 & T_{MAXIT} with N parent string.
6. For each parent i , generate $m(i)$ children using crossover
7. Find the best child for each parent (1st level   of competition).
8. Select the best child as the parent for the next generation. For each family accept the best child as the parent for the next generation if

$$Y_1\ <\ Y_2\ \ OR\ \ exp[(Y_2 - Y_1)\ /\ T]\ \geq \rho$$

   where
   Y_1 is the objective value of the best child
   Y_2 is the objective value of its parent
   T is the temperature co-efficient
   ρ is a random number uniformly distributed between 0 and 1.
9. Repeat step 10 to Step 13 for each family
10. Count = 0
11. Repeat step 11 for each child: Go to step13
12. Increase count by 1, if
   $((Y_1<Y_2)\ exp\ ((Y_{LOWEST}-Y_1)/T) \geq \rho)$
   where Y_1 is the objective value of the child
   $Y2$ is the objective value of its parent
   Y_{LOWEST} is the lowest objective value ever found
   T is the current temperature
   ρ is a random number uniformly distributed between 0 and 1.
13. Acceptance number of the family is equal to count (A)
14. Sum up the acceptance number of all the families (S)
15. For each family i, calculate the number of children to be generated in the next generation according to the following formula $m\ (i) = (TC \times A)\ /\ S$
   where, TC is the total number of children generated by all the families.
16. Decrease the temperature.
17. Repeat Step 6 to Step 16 until a certain number of iterations has been reached.

IV. NEURO FUZZY MODELING OF END MILLING PROCESS

Intelligent manufacturing systems require intelligent models that can help the manufacturer to meet the customer demands with existing infrastructure. The rising demand for precision and quality in manufacturing necessitates that vast amounts of manufacturing knowledge be incorporated in manufacturing systems. Neuro Fuzzy modeling of end milling process is attempted in this section. Surface finish in end milling depends upon a number of variables such as cutting speed, feed rate, spindle speed, radial depth of cut, tolerance etc. The relative effect of these variables on surface roughness and machining time is quite considerable. A complex relationship exists between these process parameters and hence there is a need to develop intelligent models which can capture this complex interrelationship and enable fast computation of the average surface roughness and machining time based on process parameters.

Tansel et al. [6] have experimentally measured the surface roughness and the machining time at various test conditions. Aluminum block having 30x30x90mm dimensions was machined at three stages. The first two stages, rough and semi finish cut were the same for the entire part. A flat end mill with a 12mm diameter was used for rough cutting. The

depth of cut was 1.5mm. 3D spiral tool motions were performed with 3mm stopovers at 2500mm/min feed rate and 5000rpm spindle speed. The rough cutting continued until 0.6 mm thick material was left on the desired final surface. A ball end mill with a 12mm diameter was used at the second stage to machine the material with a 0.3mm depth of cut. The step over, feed rate and spindle speed were 3mm, 700mm/min, and 3000rpm, respectively. After the second stage, 0.3 mm thick material was left on the desired part surface. The finishing cut (Third stage) was performed with a ball end mill with 10mm diameter.

Finishing cut continued until the desired surface was obtained. The surface roughness of the machined surface was measured by using a Mitatoyu Surftest 301 portable surface roughness tester. The surface roughness was measured three times at 10 different regions for each cutting condition and average was calculated.

The ranges of the cutting parameters are presented in table 1 and the sample experimental values by Tansel et al. and estimated values by neuro fuzzy model are shown in table 2.

Cutting Speed (m/min)	Feed Rate (mm/tooth)	Radial Depth of cut (mm)	Tolerance (mm)
74-123	0.07-0.12	0.1-0.3	0.01-0.001

Table 1 Range of cutting parameters

No	Cutting Speed (m/min)	Feed rate (mm/tooth)	Radial depth of cut (mm)	Tolerance (mm)	Experimental values by Tansel et al. [6]		Estimated values using Neuro Fuzzy Model	
					Average surface roughness (μm)	Machining Time (min)	Average surface roughness (μm)	Machining Time (min)
1	74	0.07	0.1	0.001	0.26	64	0.26	64.00
2	98.5	0.07	0.1	0.001	0.31	47	0.31	47.00
3	123	0.07	0.1	0.001	0.27	45	0.27	45.00
4	74	0.095	0.1	0.001	0.32	48	0.32	48.00
5	98.5	0.095	0.1	0.001	0.36	36	0.35	35.99
6	123	0.095	0.1	0.001	0.85	29	0.85	29.00
7	74	0.12	0.1	0.001	0.48	39	0.48	39.00
8	98.5	0.12	0.1	0.001	0.37	28	0.37	27.99
9	123	0.12	0.1	0.001	1.58	24	1.58	23.99
10	74	0.07	0.2	0.001	0.36	32	0.35	31.99
11	98.5	0.07	0.2	0.001	0.59	24	0.59	23.99
12	123	0.07	0.2	0.001	0.52	19	0.52	18.99
13	74	0.095	0.2	0.001	0.51	23	0.51	23.00
14	98.5	0.095	0.2	0.001	0.53	17	0.53	17.00
15	123	0.095	0.2	0.001	0.81	15	0.81	15.00

Table 2 Sample experimental and estimated values for given average surface roughness and machining time

Neuro-fuzzy inference system under consideration has four inputs as shown in figures 1 and 2 viz. cutting speed, feed rate, radial depth of cut, tolerance and one output machining time and average surface roughness each. The overall output is expressed as linear combinations of the consequent parameters. The output f can be written as:

$$f = \sum_{i=1}^{81} \bar{w}_i f_i$$

$$f = \sum_{i=1}^{81} (\bar{w}_i \alpha) p_i + (\bar{w}_i \beta) q_i + (\bar{w}_i \nu) r_i + (\bar{w}_i \gamma) s_i + (\bar{w}_i) t_i$$
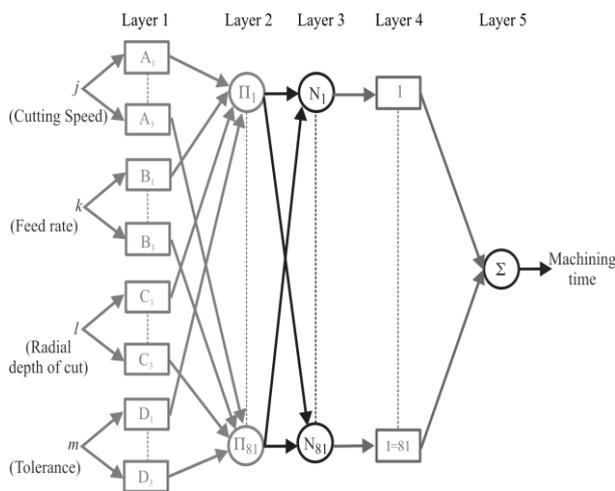

Fig. 1 A four input and one output (machining time) neuro-fuzzy network model for end milling.
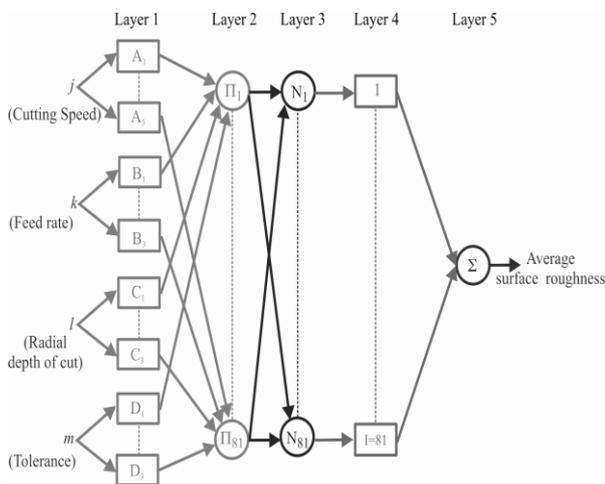
Fig. 2 A four input and one output (average surface roughness) neuro-fuzzy network model for end milling.

This is linear in the consequent parameters. The forward pass of the learning algorithm continues up to nodes at layer 4 and consequent parameters are determined by the method of least squares. In the backward pass, the error signal propagates backward to update the premise parameters by gradient descent.

The close agreement of the experimental values reported by Tansel et al. and the computed values after training NF model in table 2 clearly indicates that the model can be used for predicting the values in the range of parameters under consideration and is suitable to act as function approximator in QIET. The model is very fast and the time taken for prediction is negligible. The training information of Neuro-fuzzy model is shown in table 3.

Number of nodes:	193
Number of linear parameters:	405
Number of nonlinear parameters:	36
Total number of parameters:	441
Number of training data pairs:	81
Number of checking data pairs:	40
Number of fuzzy rules:	81

Table 3 Training parameters of NF architecture for end milling process.

V. QUANTUM INSPIRED NEURO-FUZZY EVOLUTIONARY TECHNIQUE (QINFET) AND IT'S APPLICATION TO END MILLING PROCESS

In this section end milling process is chosen to demonstrate the effectiveness of the hybrid approach formulated by integrating Neuro-Fuzzy (NF) network models, genetic algorithms (GA) and simulated annealing (SA) for process optimization to form a novel hybrid technique namely Quantum Inspired Neuro-Fuzzy Evolutionary Technique (QINFET). The optimization is performed using the Quantum Inspired Evolutionary Technique (QIET) algorithm which requires that the fitness function is easily computable for the method to be computationally tractable [7]. A NF model is used to provide the fitness function value in the QIET. Thus QINFET uses neuro-fuzzy network model in tandem with Quantum Inspired Evolutionary Technique (QIET) in determining the optimal process parameters. The NF model intelligently determines the average surface roughness and machining time for a given set of input process parameters. Once the NF model is ready it is incorporated in the QIET algorithm for fitness evaluation while finding optimal values. This integration of NF model enables fast computation of fitness function which is the primary requirement for successful implementation of the evolutionary optimization. This approach of using a Neuro-Fuzzy in QIET is quite similar to that of meta-model. The cutting conditions of end milling process were optimized using the QINFET to obtain the best compromise between two critical machining-related values: surface roughness and machining time. Spindle speed, feed

rate, radial depth of cut and tolerance were optimized, while one of the two key performance values was kept in the desired range and the other one was minimized. QINFET generated a series of alternatives for the user. The results demonstrated the compromise between the machining time and estimated surface roughness. When the minimization of the surface roughness requested, QINFET selected high cutting speed and very small feed rate. To minimize the machining time, very high cutting speed and the feed rate were selected. The surface roughness deteriorated in these cases. The tendency of the estimations of the QINFET agreed with the theoretical expectations.

The optimal parameters were found after 50 runs of QINFET algorithm for average surface roughness and machining time. Table 4 and 5 shows comparison of optimized results between those reported by Tansel et al. and QINFET algorithm.

The optimization results using QINFET as indicated in tables below show a close agreement with the optimized results reported by Tansel et al. [6] using Genetically Optimized Neural Network System (GONNS). The Quantum Inspired Neuro Fuzzy Evolutionary Technique (QINFET) is a flexible and versatile technique that can be used for intelligent modeling and optimization of process parameters.

VI. CONCLUSION

In this paper Quantum Inspired Evolutionary Technique (QIET) is presented. The technique has been carefully designed with various features that enable it to seek the near global optimum rapidly without getting stuck in the local optima. The algorithm allows a natural coding of design variables by considering continuous variables.

Range is selected for	Critical Parameters				Optimized operating conditions – the minimized critical parameter is underlined							
Machining time (min)	Machining time (min)		Surface Roughness (µm)		Cutting speed (m/min)		Feed rate (mm/tooth)		Radial depth of cut (mm)		Tolerance(mm)	
	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET
7.3-65	54.98	41.22	0.14	0.14	89.50	84.39	0.07	0.07	0.1	0.1	0.01	0.01
7.3-10	9.99	9.00	0.34	0.25	88.64	88.58	0.12	0.12	0.27	0.30	0.001	0.001
7.3-20	15.96	15.30	0.207	0.206	86.26	85.62	0.08	0.10	0.3	0.23	0.001	0.001

Table 4 Comparison of Optimization Results obtained by Tansel et al. and QINFET
(Minimization of Surface Roughness)

Range is selected for	Critical Parameters				Optimized operating conditions – the minimized critical parameter is underlined							
Surface Roughness (µm)	Machining time (min)		Surface Roughness (µm)		Cutting speed (m/min)		Feed rate (mm/tooth)		Radial depth of cut (mm)		Tolerance(mm)	
	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET	Tansel	QINFET
0.2-1.58	7.17	7.00	1.01	0.50	122.99	115.77	0.12	0.12	0.3	0.3	0.001	0.001
0.2-0.50	8.68	8.11	0.5	0.51	97.92	96.28	0.12	0.11	0.29	0.29	0.001	0.001
0.2-0.80	7.39	7.01	0.68	0.61	123	104.80	0.12	0.12	0.3	0.3	0.01	0.01

Table 5 Comparison of Optimization Results obtained by Tansel et al. and QINFET
(Minimization of Machining Time)

This may give designer more flexibility in the optimization problems. This technique is further generalized by incorporating provision to embed Neuro-Fuzzy models as fitness evaluators to create Quantum Inspired Neuro Fuzzy Evolutionary Technique (QINFET). Subsequently this new technique is applied to process optimization of end milling process and the results are presented. The proposed design scheme helps to achieve the desired level of control needed to avoid costly production problems and ensures economical production of quality products. QINFET demonstrates promise in optimizing complex industrial processes pertaining to intelligent manufacturing systems for achieving energy and material saving, quality improvement in the end product.

References

[1]   K Ganesh and M Punniyamoorthy, Optimization of continuous —time production planning using hybrid genetic algorithms—simulated annealing, *The International Journal of Advanced Manufacturing Technology,vol.26, no.1-2,* 2005, 148–154.

[2]   G Nallakumarasamy, P Srinivasan, K Venkatesh Raja, and R Malayalamurthi, Optimization of operation sequencing in CAPP using simulated annealing technique (SAT), *International Journal of Advanced Manufacturing Technology, vol. 54, no. 5-8,* 2011, 721–728.

[3]   K Hans Raj, R S Sharma, G S Mishra, A Dua, C Patvardhan, An evolutionary computational technique for constrained optimisation of engineering design, *Institution of Engineers (India) Journal, vol. 86,* 2005, 121-128.

[4]   K H Han and J H Kim, Quantum-Inspired Evolutionary Algorithm for a Class of Combinatorial Optimization, *IEEE Transactions on Evolutionary Computation, vol. 6, No. 6.*, 2002, 671-682.

[5]   K H Han and J H Kim, Quantum-inspired evolutionary algorithm with a new termination criterion, H gate, and two-phase scheme, *IEEE Transactions on Evolutionary Computation, vol. 8,* 2004, 156-169.

[6]   I N Tansel, B Ozelik , W Y Bas, P Chen, D Rincon, S Y Yang, A Yenilmeg, Selection of optimal cutting conditions by using GONNS, *International Journal of Machine Tools & Manufacture, 46,* 2006, 26-35.

[7]   K Hans Raj, S Rahul, S Rajat, A Swarup, J Rochak, A Quantum Seeded Evolutionary Computational Technique for Constrained Optimization in Engineering Design, *Proceedings in the XXXI National Systems Conference, MIT, Manipal vol.1,* 2007, 876-881.

Automotive Drum Brake Squeal Analysis Using Complex Eigenvalue Methods

Ibrahim Ahmed [1], Essam Allam[2], Mohamed Khalil[2] and Shawki Abouel-seoud[3]

*(Associate Professor, Automotive Technology Dept., Faculty of Industrial Education, Helwan University, Egypt)
** (Associate Professor, Automotive Engineering Dept., Faculty of Engineering, Helwan University, Egypt)
*** (Professor, Automotive Engineering Dept., Faculty of Engineering, Helwan University, Egypt)

ABSTRACT

Vehicle brakes can generate different kinds of noises. Eliminating brake noise is a very big challenging issue in the automotive industry. Brakes generally develop large and sustained friction-induced oscillations, referred to brake squeal. This brake squeal is considered a serious operational braking problem in passenger cars and commercial vehicles. This paper involves an approach to discover the main causes of drum squeal occurrence using finite element methods (FEM). A modal analysis of a prestressed structure will be performed to predict the onset of drum brake instability. The brake system model is based on the model information extracted from finite element models for individual brake components. An unsymmetric stiffness matrix (MATRIX27) is a result of a friction coupling between the brake lining and drum which may lead to complex eigenfrequencies. This finite element method (FEM) using ANSYS was used to predict the mode shape and natural frequency of the brake system after appropriate verification of FEM. The results showed that changing the contact stiffness of the drum-lining interface play an important role in the occurrence of the squeal. Moreover, decreasing the lining coefficient of friction lead to decreasing the occurrence of the squeal. It showed also that both the frequency separation between two systems modes due to static coupling and their associated mode shapes play an important role in mode merging. It was noted that squeals are most likely to occur when the eigenvectors and eigenvalues of the brake drum and shoes are close to the coupled vibration frequency to confirm that the coupling between different modes was necessary to form instabilities.

The results confirmed that the eigenvectors of the leading and trailing brake shoes are independent from each other with the same natural frequency.

Keywords - Drum, shoe, lining, coupling, natural frequency, instability, modal analysis and finite element.

I. INTRODUCTION

Brake squeal noise is considered a serious braking problem in passenger cars and commercial vehicles. This brake squeal noise has been researched and studied for decades and has not been fully solved yet for either drum or disc brakes. However, the squealed brake is more efficient than the non-squealed one. In particular, buses brake noise is a very serious due to that it is a major source of environmental noise pollution in big cities. Drum brake squeal noise is a complex vibration problem that has coupled vibrations, the sources of which are extremely difficult to be discovered [1]. The nature of the noise that is not repeatable at a given braking condition makes it very difficult to investigate the noise on either the real vehicle or dynamometer in terms of correlating the factors influencing the noise such as pressure, speed, and temperature.

Over the years, drum and disc brake noise have been given various names in an attempt to provide some definitions of the sound emitted such as grind, grunt, moan, groan, squeak, squeal and wire brush [2]. In general, brake noise has been divided into three categories, in relation to the frequency of noise occurrence. The three categories presented are low frequency noise, low-frequency squeal, and high-frequency squeal. Low-frequency noise of drum and disc brake is typically occurs in the frequency range

between 100 and 1000 Hz. Typical noises that reside in this category are grunt, groan, grind, and moan. This type of noise is caused by friction material excitation at the drum or rotor and lining interface. Low-frequency squeal is generally classified as a noise having a narrow frequency bandwidth in the frequency range above 1000 Hz, but below the first in-plane mode of the brake. The failure mode for this category of squeal can be associated with frictional excitation coupled with a phenomenon referred to as ''modal locking'' of brake corner components. Modal locking is the coupling of two or more modes of various structures producing optimum conditions for brake squeal [2].

High-frequency brake squeal is defined as a noise, which is produced by friction induced excitation imparted by coupled resonances (closed spaced modes) of the rotor or drums itself as well as other brake components. It is typically classified as squeal noise occurring at frequencies above 5 kHz. Since it is a range of frequency, which affects a region of high sensitivity in the human ear. High-frequency brake squeal is considered the most annoying type of brake noise. Brake squeal is a concern in the automotive industry that has challenged many researchers and engineers for years. Considerable analytical, numerical and experimental efforts have been spent on this subject, and much physical insight ha s been gained on how disc and drum brakes may generate squeal, although all the mechanisms have not been completely understood [3].

In recent years, the focus on brake squeal problems has shifted from fundamental theoretical research to more practical and problem-solving oriented efforts. Instead of a simple schematic model, the brake system model tends to include more brake components, and the effects of design parameters on the stability can be investigated. A linear system model was created based on the modal information of the brake components, and a complex eigenvalue analysis was performed to solve the Equations of motion [4]. Guan et al. [5] constructed a coupled linear model including all brake components and identified the substructure modes, which have great influence on the system stability. Chowdhary et al. [6] developed an assumed modes model for squeal prediction of a disc brake, and found that this

parathion between the frequencies is an important factor in determining the onset of flutter-type instability. Ouyang et al. [7] considered the effects of disc rotation, and the friction-induced vibration of the brake was treated as a moving load problem. With the improvement of numerical techniques, Hamabe et al. [8] and Nack [9] directly conducted a complex eigenvalue analysis with a finite element (FE) model of a brake system including the friction force. However, a nonlinear contact analysis was performed to determine the pressure distribution at the friction interface followed by system linearization and a complex modal analysis, using FE analysis [10]. Thus, in their study, the contact stiffness was dependent on local contact pressure.

The finite element method has been applied to the investigation of brake noise [11, 12], through the emphasis has been upon the investigation of drum brake noise. It can provide predictions of the vibration modes of brake components, but coupling between the brake shoe and drum for a full brake assembly model has not been fully developed. Kusano et al. [11] carried out experimental and finite element analysis to analyze the vehicle drum brake noise, using a half-brake model. Modal models representing the dynamic characteristics of the components were described with a limited number of degrees-of-freedom, then nodes on the rotor and pads were coupled through springs of the same stiffness, which were applied over the whole contact area, or normal, and friction forces were applied on the disc without modeling pads [18]. Full three-dimensional models of disc brake were used to predict the natural frequencies of the brake components [12, 13, 14], which were then connected with springs of the same stiffness. The effect of the coupling between the shoe and drum was discussed by Ghesquire and Castel [13, 14]. However, the absence of the friction force has so far been a limitation of finite element simulations of vibration while braking. Introducing friction force to the finite element model makes the stiffness matrix asymmetric. Resulting in a non-conservative system for which eigensolutions are only available in some commercially available finite element analysis programs.

A frictional counter-coupled model can illustrate the principle of drum brake vibration while squealing, as

a kind of self-excited vibration [15-16]. The squeal vibration of drum results from not some nonlinear factors such as negative slope of friction coefficient, i.e. a negative damping effect, but also the interaction between the complex mode and frictional force from sliding drum [17]. If a constant friction coefficient were large enough, the Equations of motion would have eigenvalues with positive real components, which indicates the vibration having a tendency of unlimited increase. It can also be proved by this frictional counter coupled model [15] that progressive wave is another characteristic of the squeal vibration of drum. Up till now, most of researches in the area of brake squeal depend mainly on the 2-dimensinal technique and the 3-dimensional technique is still very limited in solving the squeal problems. The main difficulty in 3-dimensional analysis of brake vibration that the movements of shoes are too complicated to be simulated in the 3-dimensional solution by Finite element. However, the mechanisms of squeal have not been understood completely yet, so that it is not easy to establish an accurate finite element model of drum brake squeal, especially in determining some boundary conditions.

The work presented here based on the conclusions in presented papers, aims to investigate the drum brake squeal noise by finite element analysis (FEA). A three dimensional elasticity dynamics model of drum brake is set up, by which the drum vibration while squealing can be investigated more accurately than 2-D model. A pressure-dependent model for the coupling between the brake drum and the shoe and lining assembly is described and used in the modal analysis of the brake assembly. Predicted vibration modes and frequencies are compared with experimental data for a drum brake noise propensity evaluation. The modal analysis of the brake assembly by FEA, using the friction interface model proposed, is found more useful for evaluation of drum brake designs for noise propensity.

II. VALIDATION OF DRUM BRAKE MODEL

An experimental modal testing is performed on the individual drum brake components to evaluate the vibration characteristics and validate the FE model. So, an impulse excitation using an impact hammer is used to obtain natural frequencies and the associated mode shape of the brake components as shown in

Fig. 1. Then, finite element models (FEM) for the brake drum and lining separately are generated as shown in Fig. 2. Each finite element model is validated by performing the modal testing on the real brake component for the free-free condition followed by modal analysis of the FE model. Free boundary conditions are used in this case because it is the easiest way to simulate experimentally.
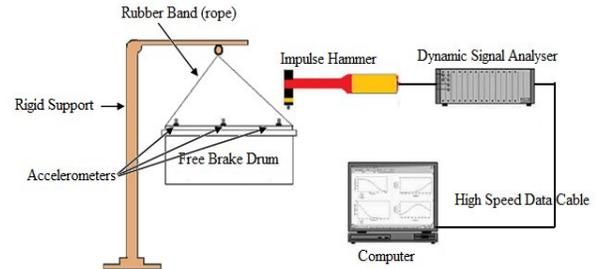
Fig. 1 Modal testing for the free-free brake drum.

As it is well known that, the drum brake assembly consists of five main parts; the drum and the two shoes and linings. The model is created directly using the ANSYS package (FEM). Fig. 2 indicates the meshed drum, shoe, and lining after appropriate simplification to the original parts. A solid45 element has been chosen from the package (ANSYS) library to model the 3-D solid structure that has 8-nodes with three degree of freedom per each node. The drum consists of 11774 brick elements with 18123 nodes, however, the shoe consists of 1260 brick elements with 2666 nodes, and lining contains 2700 brick elements with 2886 nodes. Each lining covers an angle equal to 120º of the drum ring that has an internal diameter of 340 mm as shown in Fig. 3. However, the shoe and shoe rib cover an angle equal to 140º of drum ring that has specifications as shown in Table 1.

Drum   Lining and shoe
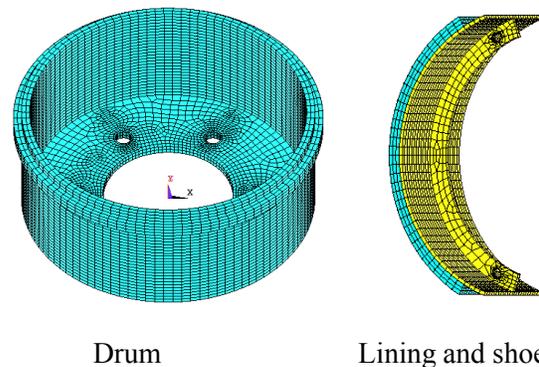
Fig. 2 FEM for the free-free brake drum and lining separately.

In the modal testing shown in Fig. 1, eight accelerometers are mounted along the drum that is hit by an impulse hummer. The tested signal is then fed from the accelerometers to a dynamic signal analyzer and finally to the computer for further analysis. The natural frequencies and modes of the drum and lining were collected in two cases. The first case was a free-free drum and free-free lining however, the second case was the coupled drum-lining with a hydraulic pressure. The collected data has been analyzed in both cases and then the finite element model is being adjusted to control the difference between the experimental results and theoretical results. These natural frequencies data have been collected in Table 1 for comparison.

Each free-free component's FE model is refined and adjusted to make the error as less as possible with experimental modal analysis results [4, 19, 20, 21, and 22]. The accurate simulation of the component models as well as the statically coupled model is important for studying the squeal characteristics either experimentally or analytically by FEM. The brake system's propensity to squeal is very sensitive to the geometry of the system and the material properties. The natural frequencies and mass-normalized mode shapes for each component are extracted from the modal analysis of the FE models. These modal characteristics of the components are used to replace the FE models to form the coupled system, and the total degrees of freedom are greatly reduced. The accuracy of the modal representations of the components and the convergence of the stability analysis results should be checked by ANSYS program.

The boundary conditions have been applied to the drum and shoe, and apply the appropriate solving to the model. Many trials were made to adjust the meshing elements for the drum and shoe with the appropriate number of elements. The lowest difference between the experimental work, and predicted FE model has been achieved in models shown in Fig. 2. These collected data showed a maximum difference of ± 3 % between the experimental and FEM results for the drum and ± 2.5 % for the shoe with lining as clear in Table 2. This difference in both cases seems to be acceptable to carry on with these models.

III. COUPLING THE DRUM-LINING FRICTION INTERFACE

As it previously said, the drum brake assembly system consists of five main parts, which are the drum and, leading shoe and trailing shoe. There is a small gap between the drum and the two shoes during the rotation of the wheel. However, this gap becomes zero at the full contact between the drum and the shoes. The coupling between the brake shoe and lining assembly and the brake drum is made by the contact between them when the brake is actuated and friction force between the lining and drum is generated as shown in Fig. 3. The coupling can be regarded as a "contact stiffness" modeled by springs connecting the brake shoe assembly and the brake drum [11].

However, the model proposed here expects that the contact force between the brake lining and brake drum will determine the degree of coupling, and thus the contact spring stiffness. This means that the contact stiffness over the whole contact area is dependent not only on the brake force applied, but also on the friction interface pressure distribution. The contact stiffness will therefore vary around the contact surface, being higher as the local contact pressure increases. The coupling between the lining and drum can be represented by springs whose stiffnesses represent the local interface contact pressure and the brake shoe can thus be modeled as being coupled to the drum via two springs, one representing the contact stiffness $K_{contact}$ and one representing the brake lining dynamic stiffness K_{lining} [1]. These two springs can then be combined to give a single "coupling" spring whose stiffness is:

$$K_{coupling} = \frac{K_{contact} \times K_{lining}}{K_{contact} + K_{lining}} \qquad (1)$$
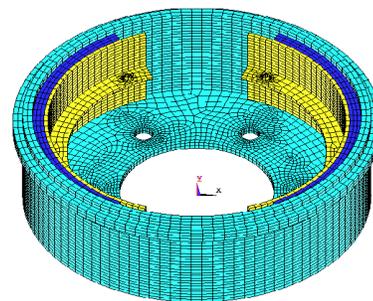

Fig. 3 FEM for the coupled drum brake system assembly.

Table 1 Original brake drum and lining properties and dimensions.

Drum dimensions and properties	Lining and shoe dimensions and properties
Outer drum diameter, 360 mm Inner drum diameter, 340 mm Inner drum cap diameter, 160 mm Outer drum cap diameter, 340 mm Height of drum cap, 10 mm Diameter of holes centerline, 220 mm Height of drum, 150 mm Holes diameter (6), 10 mm Inner diameter of upper ring, 360 mm Outer diameter of upper ring, 360 mm Height of upper ring, 15 mm Density of drum, 7350 kg/m^3 Young's modulus of drum, 1200 GN/m^2 Poisson ratio of drum, 0.27	Thickness of lining, 12 mm Width of lining, 120 mm Thickness of shoe, 4 mm Width of shoe, 120 mm Thickness of shoe rib, 4 mm Height of shoe rib, 20 mm Lining arc, 120º Shoe and rib arc, 140º Hole diameter, 10 mm Density of lining, 1350 kg/m^3 Young's modulus of lining, 200 MN/m^2 Poisson ratio of lining, 0.23 Density of shoe, 7800 kg/m^3 Young's modulus of shoe, 2000 GN/m^2 Poisson ratio of shoe, 0.27

Table 2 Natural frequencies for the modal testing and FE model for free-free brake drum and lining.

Component name	FE	No. of nodes and elements	Mode number	Modal testing (kHz).	FE model (kHz).	Difference
Drum		11774 elements and 18123 nodes	1	1483	1503	1.3 %
			2	1488	1508	1.3 %
			3	2087	2068	-1 %
			4	2140	2200	3 %
			5	2302	2332	1.3 %
			6	2340	2345	0.2 %
			7	2917	2906	-0.4 %
			8	3087	3097	0.3 %
			9	3129	3109	-0.6 %
			10	3203	3253	1.6
Lining		Lining 2700 elements and 2886 nodes + Shoe 1260 elements and 2666 nodes	1	1179	1200	2 %
			2	1213	1240	2.5 %
			3	1667	1684	1 %
			4	1843	1820	-1.2 %
			5	2226	2210	-0.7 %
			6	2233	2253	-0.9 %
			7	2605	2600	-0.2 %
			8	2679	2677	0
			9	2861	2886	1 %
			10	3109	3100	-0.3 %
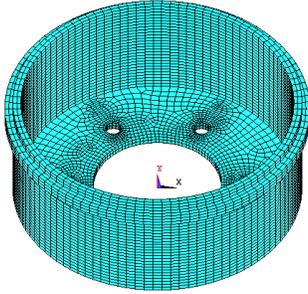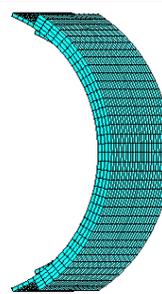
IV. MODELING OF DRUM BRAKE ASSEMBLY.

The steps of creating a linear drum brake assembly model contain three main steps as shown in Fig.4, which are:

- Pre-processor.
- Solution.
- Post-processors.

These three steps include constructing FE models for brake components by ANSYS CAD tools as mentioned previously. A modal analysis technique will be performed to extract the modal eigenvalues and eigenvectors.

Studying the effects of different parameters such as contact and lining stiffnesses and friction forces on the occurrence of squeal incorporating the effects of boundary conditions to form a coupled model.

Fig. 3 shows the finite element model for the coupled drum brake system assembly that contain the main five components which are drum, two shoes and two linings (leading and trailing), participate in the vibrational response of a drum brake system [20-21].
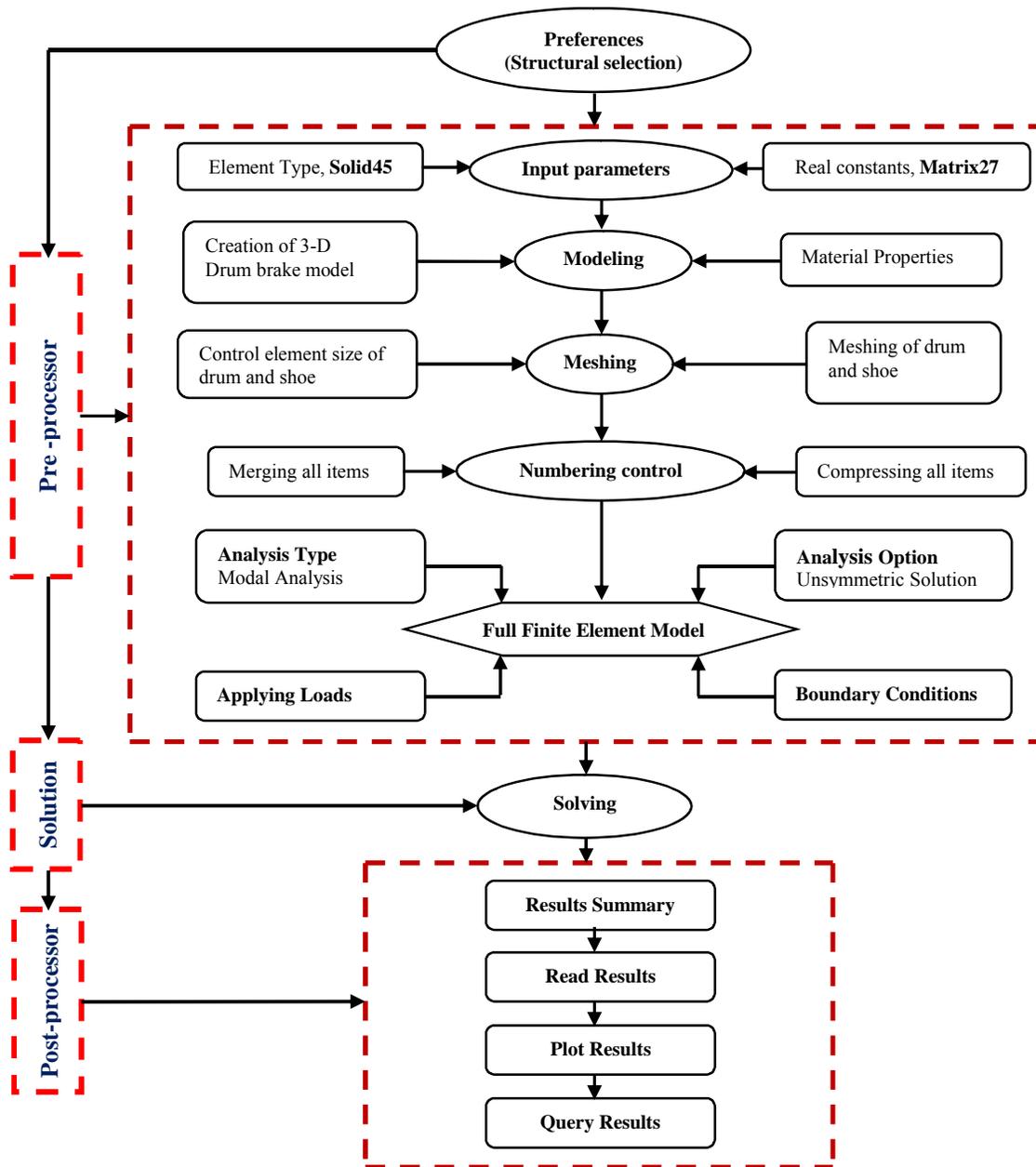

Figure 4 Parametric modeling and modal analysis flowchart.

The attached linings to the shoes will be in contact with the drum during braking to produce the friction forces. These leading and trailing shoes can be moved in different direction opposed to each other through hydraulic cylinders that contain two pistons to assist the shoes in the braking action. It is well-known that the friction-induced vibration is generated by the stiffness and friction coupling between the drum and the shoes through the shoe lining. A simplified coupled model that includes the drum, the shoes, and the shoe lining has been modelled to incorporate the effect of different boundary conditions on the occurrence of squeal as clear in Fig. 5.

Finite Element models of the drum and the shoe with lining separately are built-up by ANSYS package using 3-dimensional brick elements called SOLID45 [24] is clearly shown in Fig. 2. However, the finite element model of the coupled drum brake is shown in Fig. 3. The drum is clamped in the bolt holes (6 holes) positions while the shoe model uses free boundary conditions. The shoe lining is modeled as an integral part of the FE model of the shoe to include the inertial and stiffness influences of the shoe lining on modal characteristics (eigenvalues and eigenvectors) of the shoe.
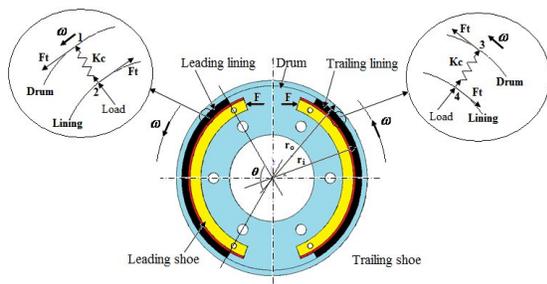
Fig. 5 Schematic diagram of the drum brake assembly showing piston pressure and contact stiffness representation.

The Equations of motion of the uncoupled system including the drum and the identical shoes can be written as [1, 20, 21 and 23]:

$$\{\ddot{q}\} + [\omega^2]\{q\} = \{0\} \qquad (2)$$

Where $[\omega^2]$ is a diagonal matrix of the extracted natural frequencies of the components, and $\{q\}$ is an N-vector of generalized coordinates. However, the number of degrees of freedom of the system, N, is equal to the total number of extracted component modes, [20].

In considering the coupling between the drum and the shoe through the contact lining, the contact interface between the drum and shoe is discretized into a mesh of 2-dimensional contact elements as shown in Fig. 5. The lining material is then modelled as a spring located at the contact elements as mentioned previously. The Equation of motion of the coupled system can be written as follows [20-21];

$$\{\ddot{q}\} + [[\omega^2] + [A] + \mu[[B] + [C]]]\{q\} = \{0\} \quad (3)$$

Where, [A] and [C] are stiffness contribution due to the lining and shoe supports respectively and [B] arises from friction coupling between the shoes and drum which is assemmetric. In the absence of lining coupling i.e., [A] and μ equal to zero, the eigenvalues are purely imaginary that is being the natural frequencies of the drum components and the shoes coupled through the hydraulic cylinder stiffness and backing plate stiffness. The solution of the Equation 3 gives the eigenvalues of:

$$S = \sigma_i \pm j\omega_i \qquad (4)$$

However; in the presence of the lining stiffness coupling but without friction coupling, the eigenvalues are again purely imaginary and correspond to the natural frequencies of an engaged brake system which is not rotating. This is referred to what is called statically coupled system.

In the presence of the lining stiffness coupling but with friction coupling and the [B] is non-symmetric. When all of the eigenvalue are purely imaginary, these correspond to the natural frequencies of an engaged and rotating system.   If any of the eigenvalues is complex, it will appear in the form of complex conjugate pairs, one with positive real part and the other with negative real part. The existence of complex roots with positive real parts indicates the presence of mode merging or what is called coupled mode, instability, which causes the brake to squeal. The value of friction coefficient that demarcates stable and unstable oscillations will be referred as a critical value of friction coefficient μ_{cr} .   The imaginary part of the eigenvalues with a doublet root

at this μ_{cr} is the squeal frequency and the corresponding mode of the complex structure is the mode shape at this squeal frequency, [20-21].

V. COMPLEX EIGENVALUE ANALYSIS.

The complex eigenvalue analysis technique that is available in ANSYS package is used to determine the stability of drum brake assembly. The real and imaginary parts of the complex eigenvalues are responsible for the degree of instability (unstable frequencies and unstable modes) of the drum brake assembly and are thought to imply the likelihood of squeal occurrence. The importance of this method lies in the asymmetric stiffness matrix that is derived from the contact stiffness and the friction coefficient at the drum-lining interface [14]. In order to perform the complex eigenvalue analysis using ANSYS, four main steps are required [15]. They are given as follows:

1. Nonlinear static analysis for applying drum brake-line pressure.
2. Nonlinear static analysis to impose rotational velocity on the drum.
3. Normal mode analysis to extract natural frequency of undamped system.
4. Complex eigenvalue analysis that incorporates the effect of friction coupling.

In this analysis, the complex eigenvalues using ANSYS are solved using the unsymmetric method. The Equation of motion of any vibrating system is:

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + [K]\{u\} = \{F\} \qquad (5)$$

Where M, C and K are mass, damping and stiffness matrices, respectively, and u is the generalized displacement vector. For friction induced vibration, it is assumed that the forcing function F is mainly contributed to by the variable friction force at the drum-lining interface. The friction interface is modelled as an array of friction springs as shown in Fig. 5. With this simplified interface model, the force vector becomes linear:

$$F = \lfloor K_f \rfloor \{u\} \qquad (6)$$

Where $\lfloor K_f \rfloor$ is the friction stiffness matrix. A homogeneous Equation is the obtained by combining

Equations 5 and 6 and by moving the friction term to the left hand side as follows:

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + \lfloor K - K_f \rfloor \{u\} = \{0\} \qquad (7)$$

Equation 7 is now the Equation of motion for a free vibration system with a pseudo forcing function in the stiffness term. The friction stiffness acts as the co-called "direct current" spring [23] that causes the stiffness matrix to be asymmetric.

The unsymmetric method, which also uses the full [K] and [M] matrices, is meant for problems where the stiffness and mass matrices are unsymmetric. It uses Lanczos algorithm, Theory Manual [24] that calculate complex eigenvalues and eigenvectors for any system. Matrix27 represented an arbitrary element whose geometry was undefined but whose elastic kinematics response could be specified by stiffness, damping, or mass coefficients. This element matrix27 was assumed to relate two nodes, each with six degrees of freedom per node: translations in the nodal x, y, and z directions and rotations about the nodal x, y, and z-axes as shown in Fig. 5. The stiffness, damping, or mass matrix constants were input as real constants. All matrices generated by this element were 12 by 12. The degrees of freedom were ordered as UX, UY, UZ, and ROTX, ROTY, ROTZ for node 1 followed by the same for node 2 on the leading shoe and also for node 3 and node 4 on the trailing shoe. If one node was not used, so all rows and columns related to this node would be defaulted to zero.

For most brake design, including this study, there was no viscous damping present. According to the geometric instability, variable frictional forces, due to variable normal forces, caused brake squeal to occur [4]. The matrix Equation 5 in the absence of viscous damping and including the frictional forces can be rewritten to be in the form of:

$$[M]\{\ddot{U}\} + [K]\{U\} = \{F_f\} \qquad (8)$$

To allow variable normal forces at the drum-lining interface, adjacent nodes on the drum and lining interface were connected together with stiff spring as shown in Figure 5. Since squeal typically occurred at low applying pressures, a constant contact pressure

was assumed between both lining material surface and the drum surface [4 and 26]. As the connected nodes moved towards or away from each other, the magnitude of the normal force increased or decreased according to moving direction. So, the resulting variable frictional forces were written in terms of relative displacement between the mating surfaces as in Equation 6. The matrix, which related the frictional forces to nodal displacement called the frictional stiffness matrix $[K_f]$ or the friction matrix.

To obtain a homogeneous Equation, the forces were moved from the right side of Equation 8 to the left side of the same Equation to be in the form of;

$$[M]\{\ddot{U}\} + [K - K_f]\{U\} = 0 \qquad (9)$$

The complementary solution to the homogeneous, second order, matrix differential Equation 9 is in the form of:

$$\{U\} = \{\phi\}e^{st} \qquad (10)$$

Where
s   is the eigenvalue.
$\{\phi\}$ is the eigenvector.

And by substitution in Equation 9 So;

$$([M]S^2 + [K - K_f])\{\phi\} = \{0\} \qquad (11)$$

The eigenvalue and possibly the eigenvectors of Equation 11 were complex numbers. Complex numbers contained two parts; real and imaginary parts. For this drum brake system, the eigenvalues always occurred in complex conjugate pairs. For certain mode, the eigenvalue was:

$$S = \sigma_i \pm j\omega_i \qquad (12)$$

where σ_i is the real part and represented the natural frequency of the system. However; ω_i is the imaginary part and represented the instability of the system.

A positive damping coefficient causes the amplitude of oscillations to increase with time. Therefore the system is not stable when the damping coefficient is positive. By examining the real part of the system eigenvalues the modes that are unstable and likely to produce squeal are revealed. In another definition of the damping ratio, which is defined as $-2\sigma/|\omega|$.

If the damping ratio is negative, the system becomes unstable, and if the damping ratio is positive, the system becomes stable.

The first unsymmetric MATRIX27 that has been generated between node 1 and node 2, will be performed as a real constant between the drum and the leading shoe as shown in Fig. 5.

$$
\begin{bmatrix} F_{X1} \\ F_{Y1} \\ F_{Z1} \\ M_{X1} \\ M_{Y1} \\ M_{Z1} \\ F_{X2} \\ F_{Y2} \\ F_{Z2} \\ M_{X2} \\ M_{Y2} \\ M_{Z2} \end{bmatrix} =
\begin{bmatrix}
0 & 0 & \mu.Sin\theta.K_C & 0 & 0 & 0 & 0 & 0 & -\mu.Sin\theta.K_C & 0 & 0 & 0 \\
0 & 0 & \mu.Cos\theta.K_C & 0 & 0 & 0 & 0 & 0 & -\mu.Cos\theta.K_C & 0 & 0 & 0 \\
0 & 0 & K_C & & 0 & 0 & 0 & 0 & -K_C & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\
0 & 0 & -\mu.Sin\theta.K_C & 0 & 0 & 0 & 0 & 0 & \mu.Sin\theta.K_C & 0 & 0 & 0 \\
0 & 0 & -\mu.Cos\theta.K_C & 0 & 0 & 0 & 0 & 0 & \mu.Cos\theta.K_C & 0 & 0 & 0 \\
0 & 0 & -K_C & & 0 & 0 & 0 & 0 & K_C & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\
0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0
\end{bmatrix}
\times
\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \theta_{X1} \\ \theta_{Y1} \\ \theta_{Z1} \\ X_2 \\ Y_2 \\ Z_2 \\ \theta_{X2} \\ \theta_{Y2} \\ \theta_{Z2} \end{bmatrix}
$$

However, the second unsymmetric MATRIX27 that has been generated between node 3 nd node 4, will be performed as a real constant between the drum and the trailing shoe as clear in Fig. 5.

$$
\begin{bmatrix}
F_{X3} \\
F_{Y3} \\
F_{Z3} \\
M_{X3} \\
M_{Y3} \\
M_{Z3} \\
F_{X4} \\
F_{Y4} \\
F_{Z4} \\
M_{X4} \\
M_{Y4} \\
M_{Z4}
\end{bmatrix}
=
\begin{bmatrix}
0 & 0 & -\mu.Sin\theta.K_C & 0 & 0 & 0 & 0 & 0 & \mu.Sin\theta.K_C & 0 & 0 & 0 \\
0 & 0 & -\mu.Cos\theta.K_C & 0 & 0 & 0 & 0 & 0 & \mu.Cos\theta.K_C & 0 & 0 & 0 \\
0 & 0 & K_C & 0 & 0 & 0 & 0 & 0 & -K_C & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mu.Sin\theta.K_C & 0 & 0 & 0 & 0 & 0 & -\mu.Sin\theta.K_C & 0 & 0 & 0 \\
0 & 0 & \mu.Cos\theta.K_C & 0 & 0 & 0 & 0 & 0 & -\mu.Cos\theta.K_C & 0 & 0 & 0 \\
0 & 0 & -K_C & 0 & 0 & 0 & 0 & 0 & K_C & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\times
\begin{bmatrix}
X_3 \\
Y_3 \\
Z_3 \\
\theta_{X3} \\
\theta_{Y3} \\
\theta_{Z3} \\
X_4 \\
Y_4 \\
Z_4 \\
\theta_{X4} \\
\theta_{Y4} \\
\theta_{Z4}
\end{bmatrix}
$$

VI. PREDICTED RESULTS OF THE FINITE ELEMENT AND DISCUSSIONS.

The results herein represent the described FEM of drum brake system as shown in Fig. 3 that its dimensions and properties are given in Table 1 and be validated as discussed previously. A total number of 100 modes were extracted through ANSYS by coupling of the drum brake items in the frequency range 1-15 kHz (squeal range). These modes include different modes of the drum and the two shoes separately as shown in Figs 6, 7, 8 and 9. These modes include six rigid body modes for each shoe, modes up to the natural frequency of 7366 Hz for the shoes, and non-rigid body modes up to frequency of 14859 Hz for the drum. However, for the coupled modes, it gives frequencies up to 5094 Hz. Figs 6 and 7 show a sample of selected drum modes and displacement contour equivalent at each mode.

The first mode occurred at 1482 Hz with 4 nodal lines as shown clearly in Fig. 6-a and five nodal lines at frequency 2302 Hz as clear in Fig. 6-b. A mix of 6, 7, 8 and 9 nodal lines appear clearly in Fig. 6-c, d, e and f that have been reached in FEM free-free drum.

Fig. 7 shows the displacement contour occurred at modes 1, 5, 12, 36, 54 and 64 respectively showing the maximum and minimum displacement that occurred on the drum ring or drum cap due to the in-plane vibrational mode.

Fig. 8 shows some selected shoe modes; which is a combination between the bending mode and twisting mode of the shoe. The modal analysis of the free-free shoe mode includes six non-rigid body modes for each shoe up to natural frequency of 7366 Hz. It showed a mixed of longitudinal bending mode and lateral bending mode as clear in the 1st mode and 11th mode at frequencies of 1179 and 3516 Hz respectively however; a combination of lateral and longitudinal twisting mode that occurred at 6th and 19th modes at frequencies of 2233 and 4383 Hz respectively. Fig. 9 shows some selected displacement contour at modes 1, 2, 6 and 38 showing a maximum and minimum displacement that occurred due to modal analysis of the free-free shoe of the drum brake.
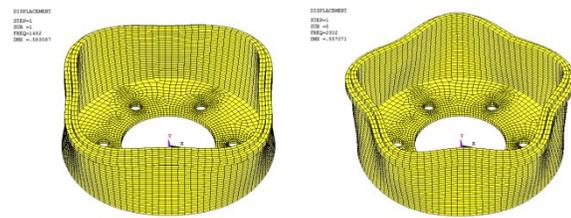
a) 1ˢᵗ mode at 1482 Hz.   b) 5ᵗʰ mode at 2302 Hz.
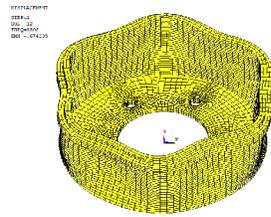
c) 12ᵗʰ mode at 3308 Hz.   d) 36ᵗʰ mode at 5969 Hz.
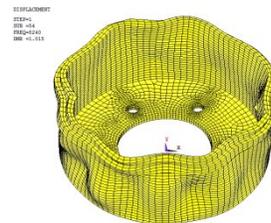
e) 54ᵗʰ mode 5240 Hz.   f) 64ᵗʰ mode at 9741 Hz.

Fig. 6 Sample of selected drum natural frequencies and mode shapes.

a) Disp. contour at 1ˢᵗ mode.   b) Disp. contour at 5ᵗʰ mode.
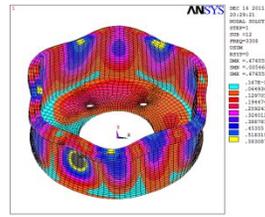
c) Disp. contour at 12ᵗʰ mode.   d) Disp. contour at 36ᵗʰ mode.
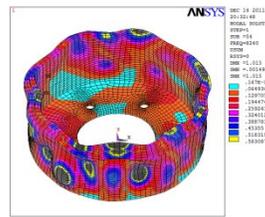
e) Disp. contour at 54ᵗʰ mode.   f) Disp. contour at 64ᵗʰ mode.

Fig. 7 Displacement contour for selected drum modes.

A total number of 100 modes were extracted through ANSYS by coupling of the drum brake items in the frequency range 1-15 kHz (squeal range) as mentioned previously. These modes include different modes of the drum and the two shoes separately and combined as shown in Fig. 10. These modes include rigid body and non-rigid body mode for both drum and two shoes up to frequencies of 5094 Hz. Fig. 10 shows a mix of mode nodal lines for the coupled drum brake ranging from 0 to 8 lines which is very clear at modes 5, 11, 14, 33, 37, 61 and 96 respectively. Coupled modes 1 and 2 do not indicate any nodal lines or circumferential circles on the drum ring surface however; it showed a shoe bending mode type at frequencies 1154 and 1218 Hz.

It showed also the in-plane vibration mode which is clear at the drum hub. The natural frequencies occurred at each component of the drum brake system at 0.42 coefficient friction. It is very clear that the drum modes occur below 5100 Hz. A number of two nodal lines appear in some modes such as 11ᵗʰ and 12ᵗʰ mode as shown in Fig. 10-d and 10-e. However; 3 nodal lines appear in the 33ʳᵈ mode and 4 nodal lines in the 5ᵗʰ mode as examples of these lines of modes. The number of nodal lines increases with the increase of natural frequency to be 5 nodal lines as shown in 14ᵗʰ mode and 6 nodal lines in the 37ᵗʰ mode. A number of 7 nodal lines appear very rare as shown in Fig. 10-i at 61ˢᵗ mode and the 8 nodal lines appears once in the 96ᵗʰ mode as also shown in Fig. 10-j.

a) **Disp. contour at 1st mode.**   b) **Disp. contour at 2nd mode.**

c) **Disp. contour at 6th mode.**   d) **Disp. contour at 38th mode.**

Fig. 9 Displacement contour for selected shoe modes.

Fig. 11 shows some selected displacement contour for the coupled drum brake system at modes 14, 17, 31, 36, 61and 96 showing a maximum and minimum displacement that occurred due to modal analysis of the coupled system. This figure indicates the vibrating parts of the coupled drum brake such as 17th mode that shows the vibrating trailing shoe. Modes 31st and 61st show the vibrating drum hub (drum cap); however the last 3 modes which are 14, 36 and 96 show the vibrating drum ring.

 Fig. 12-a and 12-b show the real part that represents the instabilities of the system when it is positive values or stabilities of the system when it is negative values, against the imaginary part which represents the natural frequencies of the system in the squeal range at 10 and 30 rad/sec of drum rotation. It is realized that as the rotation increase the system seems to be more stable than at a low speeds as clear from figures. It is also realized that the maximum natural frequency of the drum is 5094 Hz occurs at mode 100. However, the maximum natural frequency of the leading shoe is 4940 Hz occurs at mode 91 and the natural frequency of the trailing shoes is 4715 Hz occurs at mode 82 of the coupled drum brake system. The phase angle of this mode is 2º, which equal to

phase angle = $\tan^{-1}\left(\dfrac{Im.Part}{Re.Part}\right)$ [24]. When the

Fig. 8 Sample of selected shoe natural frequencies and mode shapes.

The predicted eigenvalues of the drum brake system at different drum rotating speeds of 0, 10, 20 and 30 rad/sec, whose frequencies are within the measurable frequency range , are presented in Table 3 at moderate friction coefficient of 0.42. The predicted unstable frequencies for the coupled drum brake system are 1154, 1218, 1315, 1458 and 1748 Hz which are the first five natural frequencies.

two real parts are the same with different signs, it means that they have a phase difference of 180º. It can be clear that the squealing frequency for the leading and trailing occur alternatively between each other. It is very apparent in Figs. 10-a & 10-b that the coupled system is slow below 5 kHz.

One of the main objectives of this study is to understand the effect of mode merging on the occurrence of brake squeal noise, which is not dependant on the kind of the used brake. The eigenvalue analysis of the drum coupling shown in Equation 3 gives a complex mode that contain two parts, the first part is the real part (instability) and the other part is the imaginary part (natural frequency) as mentioned earlier. The solution of this Equation achieved by substitution with the drum, lining and shoe properties shown in Table 1. Fig.13 shows the variation of the natural frequencies with changing the coefficient of friction in the range (0- 0.8) at the leading and trailing contact stiffness's. It is known that the eigenvalues and eigenvectors change when friction coefficient varies. At the critical value of friction coefficient μ_{cr} , a pair of modes merge, i.e., their frequencies and mode shapes become identical. This merging condition describes the onset of squeal very clearly [25], as seen in Fig. 13.

a) 1st mode at 1154 Hz.

b) 2nd mode at 1218 Hz.

c) 5th mode at 1458 Hz.

d) 11th mode at 1984 Hz.

e) 12th mode at 1984 Hz.

f) 14th mode at 2222 Hz.

g) 33rd mode at 2963 Hz.

h) 37th mode at 3087 Hz.

i) 61st mode at 4027 Hz.

j) 96th mode at 5046 Hz.

Fig. 10 Sample of selected coupled mode frequencies and modes shapes.

a) Disp. contour at 14th mode.

b) Disp. contour at 17th mode.

c) Disp. contour at 31st mode.

d) Disp. contour at 36th mode.

e) Disp. contour at 61st mode.

f) Disp. contour at 96th mode.

Fig. 11 Displacement contour for selected coupled modes.

The mode-merging instability due to leading-drum coupling occurs four times during this coupling in the frequency range 1-15 kHz at friction coefficients of (0-0.8) and leading contact stiffness of 200 MN/m as shown in Fig. 13-a. The first mode merging occurs between modes 1 & 2 at μ= 0.22 with a torsional mode shape and this friction coefficient called critical friction coefficient μ_{cr} .

bending mode shapes and a natural frequency near to 4150 Hz. The last mode merging occurs between 94th and 95th modes at $\mu_{cr} = 0.71$ with a combination of torsional and bending mode shape and a natural frequency close to 5310 Hz.

Table 3 The first five predicted eigenvalues (Hz) of the coupled drum brake.

ω rad/sec	S_1	S_2	S_3	S_4	S_5
0	-25 ± 1154	15 ± 1218	-05 ± 1315	0.1 ± 1458	-12 ± 1748
10	-18 ± 1154	-17 ± 1218	1.4 ± 1315	7 ± 1458	1.3 ± 1748
20	2 ± 1154	5 ± 1218	-16 ± 1315	-11 ± 1458	-22 ± 1748
30	-20 ± 1154	-16 ± 1218	-5 ± 1315	0.5 ± 1458	2 ± 1748

It is realized that when μ is bigger than μ_{cr} , the corresponding eigenvalues become complex, its imaginary parts (squeal frequency) near 1218 Hz. The second mode merging occurs between modes 12 and 13 near $\mu_{cr} = 0.58$, with a bending mode shapes and a natural frequency near to 2400 Hz. The third mode merging occurs between 55th and 56th modes near $\mu_{cr} = 0.61$, with a bending mode shapes and a natural frequency near to 4100 Hz. The last mode merging occurs between 94th and 95th modes at $\mu_{cr} = 0.34$ with a combination of torsional and bending mode shape and a natural frequency close to 5094 Hz.

However, The mode-merging instability due to trailing-drum coupling occurs also four times during this coupling in the frequency range 1-15 kHz at friction coefficients (0-0.8) and trailing contact stiffness of 150 MN/m as shown in Fig.13-b. The first mode merging occurs between modes 1 and 2 at μ = 0.41 with a torsional mode shape and this friction coefficient called critical friction coefficient μ_{cr} .

The corresponding eigenvalues at this merging become complex, its imaginary parts (squeal frequency) near 1225 Hz. The second mode merging occurs between modes 12 and 13 near $\mu_{cr} = 0.15$, with a bending mode shapes and a natural frequency near to 2500 Hz. The third mode merging occurs between 55th and 56th modes near $\mu_{cr} = 0.42$, with a

It is found that squeal frequencies are often close to natural frequencies of one or more of the components or near some natural frequencies of the statically coupled system. Since both the frequencies and mode shapes change as the brake is engaged, and further change as the friction is included, it may be difficult to identify which component modes lead to squeal. Moreover, there are many more component modes in a drum brake system, and it is difficult to explain why squeal occurs only at a few frequencies.

It is also found that the modes with the least separation due to static coupling tend to merge and become complex for higher values of friction coefficient and that agrees with the fact that usually the neighbouring modes with close frequencies arising due to components symmetry tend to merge, [26]. It can be expected that the shapes of a pair of modes also play an important role in mode merging besides the closeness of their frequencies. For each statically coupled mode for the system at μ = 0, the mode shapes of the two shoes are either identical to each other or differ by a phase angle of 180° as shown clearly in Figs. 10-a and 10-b.
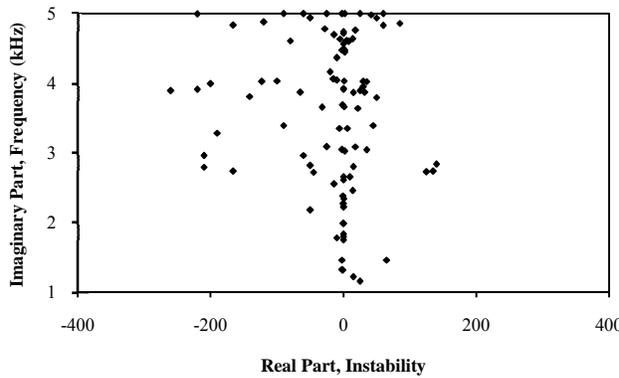
a) Rotation speed of 10 rad/sec.
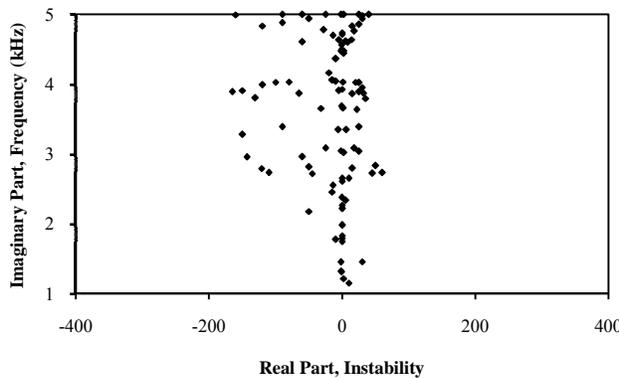
b) Rotation speed of 30 rad/sec.

Fig. 12 Real part against Imaginary part of the coupled drum brake system at 10 and 30 rad/sec.

The mode shapes of the shoes show a strong resemblance to those of the corresponding drum side wall where the two are in contact through the lining. The statically coupled modes can be divided into two groups: the first group which the two shoes move in the opposite radial direction (one moves outwards while the other moves inwards) however; the second group which the two shoes move in the same radial direction. All the modes in the same group will be termed "compatible", while the modes from different groups will be termed "incompatible" [20]. Compatible modes are more similar than incompatible modes at it $\mu = 0$ and can quite possibly become identical when μ is increased. Incompatible modes such as mode 40 and mode 55 are never seen

to combine to a merging state even though their frequency separation is quite small at $\mu = 0$.
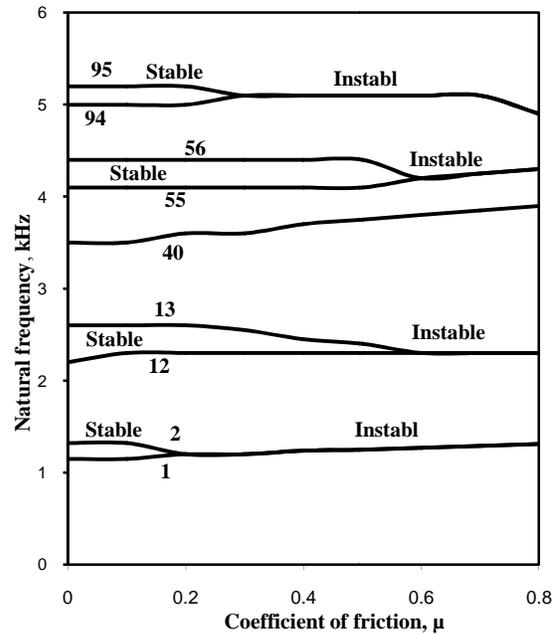
a) at leading contact stiffness of 200 MN/m.
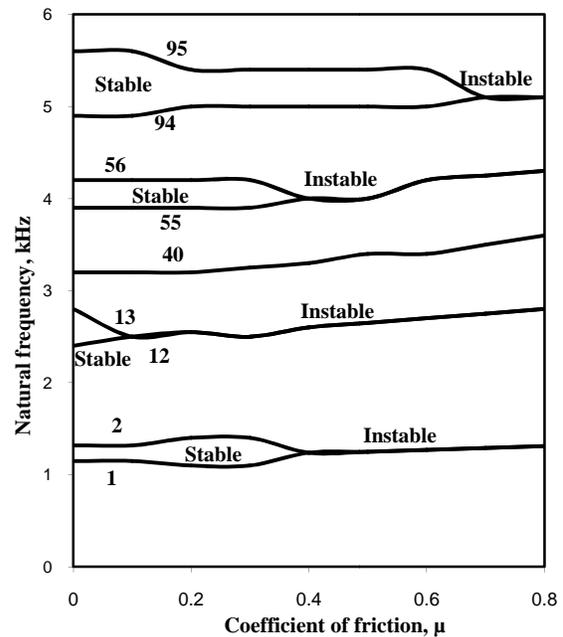
b) at trailing contact stiffness of 150 MN/m.

Fig. 13 Natural frequency of the system against the Coefficient of friction

The system can be coupled dynamically by introducing the matrices [B] and [C] shown in Equation 3. These two matrices introduce friction into the system. With the coupled system defined, it is possible to perform stability analyses to determine the conditions on system parameters that lead the brake system into squeal. A standard method of presenting stability is to plot the imaginary components of the eigenvalues (the natural frequencies) as a function of the coefficient of friction at the interface as shown in Figs.13-a, b. When a coupled system becomes unstable, it appears that the strain in the lining is minimized for the unstable modal pair of the drum and shoe. This allows for the greatest energy transfer between the shoe and drum. When the modes are nearly identical the friction force is applying the maximum force possible onto the opposite member. This results in the modal coupling, due to the curved shape of the shoes, being at its maximum value. At the same time, the frictional moment being produced is also at its maximum, thus resulting in a self-excited unstable system.

In terms of noise index (NI), an automotive drum brake system may possess many unstable vibration modes in the audible frequency range. However, not all of them result in squeal. Vibrational modes slightly unstable in the theoretical sense may never become unstable in reality due to dissipative damping in a real drum brake system. To compare the squeal propensity among unstable vibrational modes, the magnitude of the instability has been traditionally employed as a noise index. However, because the instability measurement associated with the high frequency modes are often greater than those of low frequency modes, using the magnitude of the instability as a noise index often implies high frequency squeal is more likely to occur than low frequency squeal. In this study, the noise index was defined as Yuan [28] for each vibration mode;

$$\text{Noise Index} = \frac{\sigma_j}{\sqrt{\sigma_j^2 + \omega_j^2}} \times 100 \ \% \quad (j=1,2,\dots) \ (13)$$

Where, σ_j is the instability of the system (real part) and ω_j is the natural frequency of the system (imaginary part).

The greater the noise index the more likely the corresponding mode was considered to cause audible squeal noise. It is realized in Figs. 14-a and 14-b that the maximum noise index was 5.4%, which was at frequency 2457 Hz occurred at rotating velocity of 10 rad/sec. However; themaximum noise index of 3.5% occurred at frequency of 3352 Hz at 30 rad/sec. The highest noise indexes were found at the maximum instabilities to confirm that the audible squeal noise could occur at the maximum instabilities as also indicated by Yuan [28].
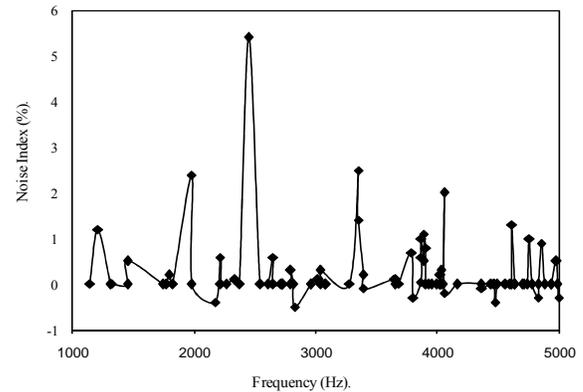
a) Noise index at rotating velocity of 10 rad/sec.
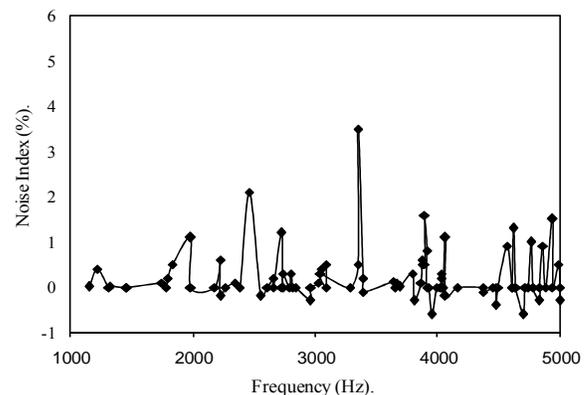
b) Noise index at rotating velocity of 30 rad/sec.

Fig. 14 the noise Index against the frequency of the coupled drum brake system.

VII. CONCLUSION

The main conclusion from this work can be summarized as follows:

- Brake squeal is a phenomenon of self-excited friction induced vibrations resulting from mode coupling. The eigenvalues and eigenvectors of a coupled drum brake system are able to provide relevant information about the modes involved in brake squeal.

- The modal behaviors components can be extracted from FEM representation of the drum brake system analysis. FE analysis makes it easy to capture geometry complexities of the components and incorporate the results of contact analysis in the system model.

- When the separation between the two modes due to static coupling is small and their mode shapes are compatible, the two modes can merge when the friction is introduced and increased. The instabilities come from the compatible modes when they are identical. So, the mode shapes of brake components most likely are measured with experimental methods has a great on mode merging. The understanding of the important role of mode shapes is expected to be of great help for the prediction of the occurrence of squeal.

- The stability boundaries are sensitive to changes in parameters such as contact stiffness.

- The changes in separations partially reveal the effects of the parameters on system stability and can provide an explanation to some squeal reduction techniques due to the correlation between μ_{cr} values and the separations of the statically coupled frequencies.

- When there is more than one mode with positive real parts of their eigenvalues, the one with the largest real part for the current set of system parameters will be the mode which drives the squeal response. The value of the coefficient of friction at which the dominant mode's eigenvalues merge and become unstable is called the critical value of the coefficient of friction.

- The greater the noise index (NI) the more likely the corresponding mode was considered to cause audible squeal noise.

REFERENCES

[1] A.J. Day, and S.Y. Kim, Noise and vibration analysis of an S-cam drum brake, Proceedings of Inst. of Mech. Engrs., Vole 210, Part D *Journal of Automobile Engineering*, pp. 35-43, IMechE 1996.

[2] K.B. Dunlap, M.A. Riehle, and R.E. Longhouse, An investigative overview of automotive disc brake noise, *SAE Paper* 1999-01-0142.

[3] F. Chen, J. Chern, and J. Swayze, Modal coupling and its effect on brake squeal, *SAE Paper* 2002-01-0922.

[4] G.D. Liles, Analysis of disc brake squeal using finite element methods, *SAE Paper* No. 891150, 1989.

[5] D. Guan, and D. Jiang, A study on disc brake squeal using finite element methods, *SAE Paper* No. 980597, 1998.

[6] H.V. Chowdhary, A.K. Bajaj, and C.M. Krousgrill, An analytical approach to model disc brake system for squeal prediction, *Proceedings of DETC 2001/VIB-21560, ASME*, Pittsburgh, PA, 2001, pp. 1–10.

[7] H. Ouyang, Q. Cao, J.E. Mottershead, and T. Treyde, Vibration and squeal of a disc brake: modelling and experimental results, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 217 (10) (2003) 867–875.

[8] T. Hamabe, I. Yamazaki, K. Yamada, H. Matsui, S. Nakagawa, and M. Kawamura, Study of a method for reducing drum brake squeal, *SAE Paper* No. 1999-01-0144, 1999.

[9] W.V. Nack, Brake squeal analysis by finite element, *International Journal of Vehicle Design* 23 (3/4) (2000) 263–275.

[10] Y.S. Lee, P.C. Brooks, D.C. Barton, and D.A. Crolla, A predictive tool to evaluate disc brake squeal propensity, Part 1: the model philosophy and the contact problem, *International Journal of Vehicle Design* 31 (3) (2003) 289–308.

[11] M. Kusano, H. Ishiduu, , S. Matsummura, and S. Washizu, Expermental study on the reduction of drum brake noise, *SAE paper* 851465, 1985.

[12] G. D. Liles, Analysis of disc brake squeal using the finite element method, *SAE paper* 891150, 1989.

[13] H. Ghesquire, and L. Castel, High frequency vibrational coupling between automobile brake disc and pads, *IMechE Autotech 1991*, seminar, paper C/427/11/021.

[14] H. Ghesquire, Brake squeal noise analysis and prediction, *IMechE Autotech 1992*, seminar, paper C389/257.

[15] J. Hulten, Some drum brake squeal mechanisms, *SAE* 951280, 1995.

[16] H.Takata, W.Jiang, and T. Nishi, The theoretical and experimental investigations on drum brake squeal, *Noise-Con 98*, April 5-8, 1998, Ypsilanti, Michigan, USA, 1998.

[17] G. Adams, Self-excited oscillations in sliding with a constant friction coefficient, DE- Vol. 84-1, 1995 *Design Engineering Technical Conference*, Volume 3- Part A, ASME, 1995.

[18] J.E. Mottershead, and S.N. Chan, Brake squeal – an analysis of symmetry and flutter instability, Friction – induced vibration, chatter, squeal and chaos, *ASME* 1992, vol. 49, pp. 87-97.

[19] I.L.M. Ahmed, *Study of the Behavior of the Vehicle Disc Brakes*, Ph.D. Thesis, University of Northumbria at Newcastle Upon Tyne, UK, 2002.

[20] J. Huang, C.M. Krousgrill, and A.K. Bjaj, Modelling of automotive drum brakes for squeal and parameter sensitivity analysis, *Journal of Sound and Vibration*, 289 (2006), 245-263, 2006.

[21] B.K. Servis, *The onset of squeal vibrations in drum brake systems resulting from a coupled mode instability*, Ph.D. Thesis, Purdue University, West Lafayette, IN, 2000.

[22] J.D. Fieldhouse, and Peter Newcomb, An Investigation Into Disc Brake Squeal Using Holographic Interferometry, *3rd International EAEC Paper* No.91084, Strasborg, June 1991.

[23] J.M. Lee, S.W. Yoo, , J.H. Kim, and C.G. Ahin, A study on the squeal of a drum brake which has shoes of non-uniform cross section, *Journal of Sound and Vibration* (2001), 240(5), 789-808, 2001.

[24] *ANSYS* user's Manual.

[25] *ABAQUS* user's manual.

[26] Y.S. Lee, P.C. Brooks, D.C. Barton, and D.A. Crolla, A Study of Disc Brake Squeal Propensity Using a Parametric Finite Element model, *European Conference on Vehicle Noise and Vibration*, IMechE, C521/009/98, 1998.

[27] A.J. Day, and P.R.J. Harding, Performance variation of cam operated drum brakes, *Proceedings of the IMechE Conference on braking of road vehicles*, 1983, paper C10/83. pp. 69-77.

[28] Yongbin Yuan, A Study of the Effect of Negative Friction-Speed Slope on Brake Squeal, *ASME Design Engineering Technical Conference* Volume 3- Part A, pp. 1153-1162, 1995.

SOME IMPUTATION METHODS IN DOUBLE SAMPLING SCHEME FOR ESTIMATION OF POPULATION MEAN

Narendra Singh Thakur, Kalpana Yadav
Centre for Mathematical Sciences (CMS), Banasthali University, Rajasthan-304022 INDIA
and
Sharad Pathak
Department of Mathematics and Statistics, Dr. H. S. Gour Central Univesity, Sagar (M.P.) INDIA

ABSTRACT

To estimate the population mean with imputation, i.e. the technique of substituting missing data, there are a number of techniques available in literature like Ratio method of imputation, Compromised method of imputation, Mean method of imputation, Ahmed's methods of imputation, F-T methods of imputation and so on. If population mean of auxiliary information is unknown then these methods are not useful and the two-phase sampling is used to obtain the population mean. This paper presents some imputation methods of for missing values in two-phase sampling. Two different sampling designs in two-phase sampling are compared under imputed data. The bias and m.s.e of suggested estimators are derived in the form of population parameters using the concept of large sample approximation. Numerical study is performed over two populations using the expressions of bias and m.s.e and efficiency compared with Ahmed's estimators.

Keywords: Estimation, Missing data, Bias, Mean squared error (M.S.E), Two-phase sampling, SRSWOR, Large sample approximation.

1. INTRODUCTION:

To overcome the problem of missing observations or non-response in sample surveys, the technique of imputation is frequently used to replace the missing data. To deal with missing values effectively Kalton et al. (1981) and Sande (1979) suggested imputation that make an incomplete data set structurally complete and its analysis simple. Imputation may also be carried out with the aid of an auxiliary variate if it is available. For example Lee et al. (1994, 1995) used the information on an auxiliary variate for the purpose of imputation. Later Singh and Horn (2000) suggested a compromised method of imputation. Ahmed et al. (2006) suggested several new imputation based estimators that use the information on an auxiliary variate and compared their performances with the mean method of imputation. Shukla (2002) discussed F-T estimator under two-phase sampling and Shukla and Thakur (2008) have proposed estimation of mean with imputation of missing data using F-T estimator. Shukla et al. (2009) have discussed on utilization of non-response auxiliary population mean in imputation for missing observations and Shukla et al. (2009a) have discussed on estimation of mean under imputation of missing data using factor type estimator in two-phase sampling. Shukla et al. (2011) suggested linear combination based imputation method for missing data in sample. The objective of the present research work is to derive some imputation method for mean estimation in case population parameter of auxiliary information is unknown.

2. NOTATIONS:

Let $U = (U_1, U_2, U_3,…,U_N)$ be the finite population of size N and the character under study be denoted by y. A large preliminary simple random sample (without replacement) S' of n' units is drawn from the population on U and a secondary sample S of size n ($n < n'$) is drawn in either two ways: One is as a sub-sample from sample S' (denoted by design I) as in fig. 1 and other is independent to sample S' (denoted by design II) as in fig. 2 without replacing S'. The sample S can be divided into two non-overlapping sub groups, the set of responding units, by R, and that of non- responding units by R^c and the number of responding units out of sampled n units be denoted by r ($r<n$). For every unit $i \in R$ the value y_i is observed, but for the units $i \in R^c$, the y_i are missing and instead imputed

values are derived. The ith value x_i of auxiliary variate is used as a source of imputation for missing data when $i \in R^C$. Assume for S, the data $x_s = \{x_i : i \in S\}$ and for $i' \in S'$, the data $\{x_{i'} : i' \in S'\}$ are known with mean $\bar{x} = (n)^{-1} \sum_{i=1}^{n} x_i$ and $\bar{x}' = (n')^{-1} \sum_{i=1}^{n'} x_i$ respectively. The following symbols are used hereafter:

$\overline{X}$, $\overline{Y}$: the population mean of X and Y respectively; $\bar{x}$, $\bar{y}$: the sample mean of X and Y respectively;

$\bar{x}_r$, $\bar{y}_r$: the sample mean of X and Y respectively; ρ_{XY} : the correlation coefficient between X and Y ;

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; $\delta_1 = \left(\frac{1}{r} - \frac{1}{n}\right)$; $\delta_2 = \left(\frac{1}{n} - \frac{1}{n'}\right)$; $\delta_3 = \left(\frac{1}{n'} - \frac{1}{N}\right)$; $\delta_4 = \left(\frac{1}{r} - \frac{1}{N-n'}\right)$; $\delta_5 = \left(\frac{1}{n} - \frac{1}{N-n'}\right)$; $f_1 = \frac{r}{n}$,

$C = \frac{(\delta_9 - \delta_4)(\delta_3 + \delta_5)}{[\delta_{10}(\delta_3 + \delta_5) - \delta_5^2]}$; $D = \frac{(\delta_{11} - \delta_4)(\delta_3 + \delta_4)}{[\delta_{11}(\delta_3 + \delta_4) - \delta_4^2]}$.

## 3. LARGE SAMPLE APPROXIMATIONS:
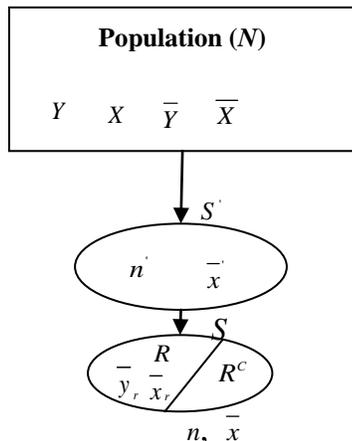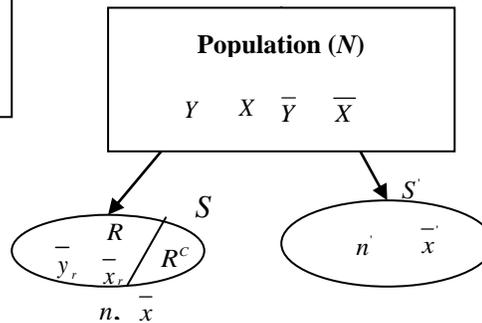
Fig. 1 [Design I]   Fig. 2 [Design II]

Let $\bar{y}_r = \overline{Y}(1 + e_1)$; $\bar{x}_r = \overline{X}(1 + e_2)$; $\bar{x} = \overline{X}(1 + e_3)$ and $\bar{x}' = \overline{X}(1 + e_3')$, which implies the results $e_1 = \frac{\bar{y}_r}{\overline{Y}} - 1$;

$e_2 = \frac{\bar{x}_r}{\overline{X}} - 1$; $e_3 = \frac{\bar{x}}{\overline{X}} - 1$ and $e_3' = \frac{\bar{x}'}{\overline{X}} - 1$. Now by using the concept of two-phase sampling and the the mechanism of MCAR, for given r, n and n' (see Rao and Sitter (1995)) we have:

Designs	$E(e_i)$	$E(e_3')$	$E(e_1^2)$	$E(e_2^2)$	$E(e_3^2)$	$E(e_3'^2)$
I	0	0	$\delta_1 C_Y^2$	$\delta_1 C_X^2$	$\delta_2 C_X^2$	$\delta_3 C_X^2$
II	0	0	$\delta_4 C_Y^2$	$\delta_4 C_X^2$	$\delta_5 C_X^2$	$\delta_3 C_X^2$

Designs	$E(e_1 e_2)$	$E(e_1 e_3)$	$E(e_1 e_3')$	$E(e_2 e_3)$	$E(e_2 e_3')$	$E(e_3 e_3')$
I	$\delta_1 \rho\, C_Y C_X$	$\delta_2 \rho\, C_Y C_X$	$\delta_3 \rho C_Y C_X$	$\delta_2 C_X^2$	$\delta_3 C_X^2$	$\delta_3 C_X^2$
II	$\delta_4 \rho\, C_Y C_X$	$\delta_5 \rho\, C_Y C_X$	0	$\delta_5 C_X^2$	0	0

4. PROPOSED STRATEGIES:

Let y'_{ji} denotes the ith observation of the jth suggested imputation strategy and b_1, b_2, b_3 are constants such that the variance of obtained estimators of $\overline{Y}$ is minimum. We suggest the following tools of imputation:

(1)
$$y'_{4i} = \begin{cases} y_i & \text{if} \quad i \in R \\[2mm] \left[\overline{y}_r + b_1\left(x_i - \overline{x}_r\right)\right] & \text{if} \quad i \in R^C \end{cases}$$
…(4.1)

under this strategy, the point estimator of $\overline{Y}$ is given by $\quad t'_4 = \overline{y}_r + b_1\left(\overline{x} - \overline{x}_r\right)$ …(4.2)

(2)
$$y'_{5i} = \begin{cases} y_i & \text{if} \quad i \in R \\[2mm] \overline{y}_r + \dfrac{b_2}{(1-f_1)}\left(\overline{x}' - \overline{x}\right) & \text{if} \quad i \in R^C \end{cases}$$
…(4.3)

under this , the estimator of $\overline{Y}$ is $\quad t'_5 = \overline{y}_r + b_2\left(\overline{x}' - \overline{x}\right)$ …(4.4)

(3)
$$y'_{6i} = \begin{cases} y_i & \text{if} \quad i \in R \\[2mm] \overline{y}_r + \dfrac{b_3}{(1-f_1)}\left(\overline{x}' - \overline{x}_r\right) & \text{if} \quad i \in R^C \end{cases}$$
…(4.5)

hence the estimator of $\overline{Y}$ is $\quad t'_6 = \overline{y}_r + b_3\left(\overline{x}' - \overline{x}_r\right)$ …(4.6)

5. BIAS AND MEAN SQUARED ERROR OF PROPOSED ESTIMATORS:

Let $B(.)_t$ and $M(.)_t$ denote the bias and mean squared error (*M.S.E.*) of an estimator under a given sampling design $t = I, II$, then the bias and m.s.e of t'_4, t'_5 and t'_6 are derived in the following theorems. The proofs of all these results are similar and therefore we will proof only one of them i.e. theorem 5.1.

Theorem 5.1:

(1) Estimator t'_4 in terms of e_i; $i = 1,2,3$ and e'_3 could be expressed:
$$t'_4 = \overline{Y}(1 + e_1) + b_1\overline{X}(e_3 - e_2)$$ …(5.1)

by ignoring the terms $E\left[e_i^r e_j^s\right]$, $E\left[e_i^r \left(e'_j\right)^s\right]$ for $r + s > 2$, where $r, s = 0,1,2,...$ and $i = 1,2,3; j = 2,3$ which is first order of approximation.

Proof: $\quad t'_4 = \overline{y}_r + b_1\left(\overline{x} - \overline{x}_r\right) \quad = \overline{Y}(1 + e_1) + b_1\overline{X}(e_3 - e_2)$

(2) Bias of t'_4 under design *I* and *II* is:

   (i) $\quad B\left[t'_4\right]_I = 0$ …(5.2)

   (ii) $\quad B\left[t'_4\right]_{II} = 0$ …(5.3)

Proof:

   (i) $\quad B\left(t'_4\right)_I = E\left[t'_4 - \overline{Y}\right]_I = \overline{Y} - \overline{Y} = 0$

   (ii) $\quad B\left(t'_4\right)_{II} = E\left[t'_4 - \overline{Y}\right]_{II} = \overline{Y} - \overline{Y} = 0$

(3) The variance of t'_4, under design *I* and *II*, upto first order of approximation could be written as:

   (i) $\quad V\left(t'_4\right)_I = \delta_1 S_Y^2 + \left(\delta_1 - \delta_2\right)\left(b_1^2 S_X^2 - 2b_1\rho S_Y S_X\right)$ …(5.4)

   (ii) $\quad V\left(t'_4\right)_I = \delta_4 S_Y^2 + \left(\delta_4 - \delta_5\right)\left(b_1^2 S_X^2 - 2b_1\rho S_Y S_X\right)$ …(5.5)

Proof: $V\left(t'_4\right) = E\left[t'_4 - \overline{Y}\right]^2 = E\left[\overline{Y}e_1 + b_1\overline{X}(e_3 - e_2)\right]^2$

$$= E\left[\overline{Y}^2 e_1^2 + b_1^2\overline{X}^2(e_3 - e_2)^2 + 2b_1\overline{Y}\overline{X}(e_3 - e_2)e_1\right]$$

$$= E\left[\overline{Y}^2 e_1^2 + b_1^2 \overline{X}^2 \left(e_3^2 + e_2^2 - 2e_2 e_3 \right) + 2b_1 \overline{Y}\,\overline{X} \left(e_1 e_3 - e_1 e_2 \right) \right] \qquad \ldots(5.6)$$

(i)   Under Design *I* (Using (5.6))

$$V(t_4')_I = \left[\; \overline{Y}^2 \delta_1 C_Y^2 + b_1^2 \overline{X}^2 \left(\delta_2 C_X^2 + \delta_1 C_X^2 - 2\delta_2 C_X^2 \right) + 2b_1 \overline{Y}\,\overline{X} \left(\delta_2 \rho C_Y C_X - \delta_1 \rho C_Y C_X \right) \; \right]$$

$$= \left[\delta_1 S_Y^2 + b_1^2 S_X^2 (\delta_1 - \delta_2) - 2b_1 (\delta_1 - \delta_2) \rho S_Y S_X \right]$$

$$= \left[\delta_1 S_Y^2 + (\delta_1 - \delta_2) \left\{ b_1^2 S_X^2 - 2b_1 \rho S_Y S_X \right\} \right]$$

(ii)   Under Design *II* (Using (5.6))

$$V(t_4')_{II} = \left[\overline{Y}^2 \delta_1 C_Y^2 + b_1^2 \overline{X}^2 \left(\delta_5 C_X^2 + \delta_4 C_X^2 - 2\delta_5 C_X^2 \right) + 2b_1 \overline{Y}\,\overline{X} \left(\delta_5 \rho C_Y C_X - \delta_4 \rho C_Y C_X \right) \right]$$

$$= \left[\delta_4 S_Y^2 + b_1^2 S_X^2 (\delta_4 - \delta_5) - 2b_1 (\delta_4 - \delta_5) \rho S_Y S_X \right]$$

(4)   The minimum variance of the t_4' is

(i) $$\left[V(t_4')_I \right]_{Min} = \left[\delta_1 - (\delta_1 - \delta_2) \rho^2 \right] S_Y^2 \quad \text{when} \quad b_1 = \rho \frac{S_Y}{S_X} \qquad \ldots(5.7)$$

(ii) $$\left[V(t_4')_{II} \right]_{Min} = \left[\delta_4 - (\delta_4 - \delta_5) \rho^2 \right] S_Y^2 \quad \text{when } b_1 = \rho \frac{S_Y}{S_X} \qquad \ldots(5.8)$$

Proof:

(i)   By differentiating (5.4) with respect to b_1 and equate to zero, we get

$$\frac{d}{db_1} \left[V(t_4')_I \right] = 0 \;\Rightarrow b_1 = \rho \frac{S_Y}{S_X}$$

After replacing the value of b_1 in (5.4), we obtained

$$\left[V(t_4')_I \right]_{Min} = \delta_1 S_Y^2 + (\delta_1 - \delta_3) \left(\rho^2 S_Y^2 - 2\rho^2 S_Y^2 \right)$$

$$= \left[\delta_1 - (\delta_1 - \delta_2) \rho^2 \right] S_Y^2$$

(ii)   Similar to (i), we proceed for (5.5), we have

$$\frac{d}{db_1} \left[V(t_4')_{II} \right] = 0 \;\Rightarrow b_1 = \rho \frac{S_Y}{S_X}$$

After replacing the value of b_1 in (5.5), we obtained

$$\left[V(t_4')_I \right]_{Min} = \delta_4 S_Y^2 + (\delta_4 - \delta_5) \left\{ \rho^2 S_Y^2 - 2\rho^2 S_Y^2 \right\}$$

$$= \left[\delta_4 - (\delta_4 - \delta_5) \rho^2 \right] S_Y^2$$

Theorem 5.2:

(5)   The estimator t_5' in terms of e_1, e_2, e_3 and e_3' is :

$$t_5' = \overline{Y}(1 + e_1) + b_2 \overline{X} \left(e_3' - e_3 \right) \qquad \ldots(5.9)$$

(6)   The bias estimator t_5', under design *I* and *II* respectively is

(i) $$B\left[t_5' \right]_I = 0 \qquad \ldots(5.10)$$

(ii) $$B\left[t_5' \right]_{II} = 0 \qquad \ldots(5.11)$$

(7)   The variance of t_5', under design *I* and *II* respectively is:

(i) $$V(t_5')_I = \delta_1 S_Y^2 + (\delta_2 - \delta_3) \left(b_2^2 S_X^2 - 2b_2 \rho S_Y S_X \right) \qquad \ldots(5.12)$$

(ii) $$V(t_5')_{II} = \delta_4 S_Y^2 + (\delta_3 + \delta_5) b_2^2 S_X^2 - 2b_2 \delta_5 \rho S_Y S_X \qquad \ldots(5.13)$$

(8)   The minimum variance of the $t_5^{'}$ is

(i) $$\left[V\!\left(t_5^{'}\right)_I\right]_{Min} = \left[\delta_1 - (\delta_2 - \delta_3)\rho^2\right] S_Y^2 \qquad \text{when} \quad b_2 = \rho\frac{S_Y}{S_X} \qquad \text{...(5.14)}$$

(ii) $$\left[V\!\left(t_5^{'}\right)_{II}\right]_{Min} = \left[\delta_4 - \delta_5^2(\delta_3 + \delta_5)^{-1}\rho^2\right] S_Y^2 \quad \text{when} \quad b_2 = \left(\frac{\delta_5}{\delta_3 + \delta_5}\right)\rho\frac{S_Y}{S_X} \qquad \text{...(5.15)}$$

Theorem 5.3:

(9)   The estimator $t_6^{'}$ in terms of e_1, e_2, e_3 and $e_3^{'}$ is :

$$t_6^{'} = \overline{Y}(1 + e_1) + b_3\,\overline{X}\!\left(e_3^{'} - e_2\right) \qquad \text{...(5.16)}$$

(10)   The bias estimator $t_6^{'}$, under design I and II respectively is:

(i) $$B\!\left[t_6^{'}\right]_I = 0 \qquad \text{...(5.17)}$$

(ii) $$B\!\left[t_6^{'}\right]_{II} = 0 \qquad \text{...(5.43)}$$

(11)   The variance of $t_6^{'}$, under F_1 and F_2 is

(i) $$V\!\left(t_6^{'}\right)_I = \delta_1 S_Y^2 + (\delta_2 - \delta_3)\!\left(b_3^2 S_X^2 - 2b_3\rho S_Y S_X\right) \qquad \text{...(5.18)}$$

(ii) $$V\!\left(t_6^{'}\right)_{II} = \delta_4 S_Y^2 + (\delta_3 + \delta_4)b_3^2 S_X^2 - 2b_3\delta_4\rho S_Y S_X \qquad \text{...(5.19)}$$

(12)   The minimum variance of the $t_6^{'}$ is

(i) $$\left[V\!\left(t_6^{'}\right)_I\right]_{Min} = \left[\delta_1 - (\delta_1 - \delta_3)\rho^2\right] S_Y^2 \qquad \text{when} \quad b_3 = \rho\frac{S_Y}{S_X} \qquad \text{...(5.20)}$$

(ii) $$\left[V\!\left(t_6^{'}\right)_{II}\right]_{Min} = \left[\delta_4 - \delta_4^2(\delta_3 + \delta_4)^{-1}\rho^2\right] S_Y^2 \quad \text{when} \quad b_3 = \left(\frac{\delta_4}{\delta_3 + \delta_4}\right)\rho\frac{S_Y}{S_X} \qquad \text{...(5.21)}$$

6. COMPARISIONS:

(1) $$\Delta_1 = \min\!\left[V(t_4)\right] - \min\!\left[V\!\left(t_4^{'}\right)_I\right] = \left[\frac{1}{n^{'}} - \frac{1}{N}\right]S_Y^2$$

$\left(t_4^{'}\right)_I$ is better than t_4,   if $\Delta_1 > 0$ $\Rightarrow \left[\dfrac{N - n^{'}}{n^{'} N}\right] > 0$ $\Rightarrow N - n^{'} > 0$ $\Rightarrow n^{'} < N$

which is always true.

(2) $$\Delta_2 = \min\!\left[V(t_4)\right] - \min\!\left[V\!\left(t_4^{'}\right)_{II}\right] = \left[\frac{1}{N - n^{'}} - \frac{1}{N}\right]S_Y^2$$

$\left(t_4^{'}\right)_{II}$ is better than t_4, if $\Delta_2 > 0$ $\Rightarrow \left[\dfrac{N - N + n^{'}}{N(N - n^{'})}\right] > 0$ $\Rightarrow n^{'} > 0$

which is always true.

(3) $$\Delta_3 = \min\!\left[V(t_5)\right] - \min\!\left[V\!\left(t_5^{'}\right)_I\right] = \left[\frac{1}{n^{'}} - \frac{1}{N}\right]S_Y^2 + \left[\frac{2}{N} - \frac{2}{n^{'}}\right]\rho^2 S_y^2$$

$\left(t_5^{'}\right)_I$ is better than t_5, if $\Delta_3 > 0$ $\Rightarrow -\dfrac{1}{2} < \rho < \dfrac{1}{2}$

(4) $$\Delta_4 = \min\!\left[V(t_5)\right] - \min\!\left[V\!\left(t_5^{'}\right)_{II}\right] = \left[\delta_9 - \delta_4\right] S_Y^2 - \left[\delta_{10} - (\delta_3 + \delta_5)^{-1}\delta_5^2\right]\rho^2 S_Y^2$$

$\left(t_5^{'}\right)_{II}$ is better than t_5, if $\Delta_4 > 0$ $\Rightarrow \rho^2 < \dfrac{(\delta_9 - \delta_4)(\delta_3 + \delta_5)}{\left[\delta_{10}(\delta_3 + \delta_5) - \delta_5^2\right]}$ $\Rightarrow -C < \rho < C$

(5) $$\Delta_5 = \min\!\left[V(t_6)\right] - \min\!\left[V\!\left(t_6^{'}\right)_I\right] = \left[\frac{1}{n^{'}} - \frac{1}{N}\right]S_Y^2 + \left[\frac{2}{N} - \frac{2}{n^{'}}\right]\rho^2 S_y^2$$

$\left(t_6^{'}\right)_{\mathrm{I}}$ is better than t_6, if $\quad \Delta_5 > 0 \quad \Rightarrow 2\left[\frac{1}{n^{'}} - \frac{1}{N}\right]\rho^2 < \left(\frac{1}{n^{'}} - \frac{1}{N}\right) \quad \Rightarrow -\frac{1}{2} < \rho < \frac{1}{2}$

(6) $\quad \Delta_6 = \min\left[V(t_6)\ \right] - \min\left[V(t_6^{'})_{II}\right] \quad = \left[\delta_{11} - \delta_4\right] S_Y^2 - \left[\delta_{11} - (\delta_3 + \delta_4)^{-1}\delta_4^2\right]\rho^2 S_Y^2$

$\left(t_6^{'}\right)_{\mathrm{II}}$ is better than t_6, if $\quad \Delta_6 > 0 \quad\quad \Rightarrow \rho^2 < \dfrac{(\delta_{11} - \delta_4)(\delta_3 + \delta_4)}{\left[\delta_{11}(\delta_3 + \delta_4) - \delta_4^2\right]} \quad\quad \Rightarrow -D < \rho < \ D$

7. NUMERICAL ILLUSTRATIONS:

We considered two populations A and B, first one is the artificial population of size $N = 200$ [source Shukla et al. (2009)] and another one is from Ahmed et al. (2006) with the following parameters:

Table 7.0 Parameters of Populations A and B

Population	N	$\overline{Y}$	$\overline{X}$	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 $n^{'} = 60$, $n = 40$, $r = 5$ for population A and $n^{'} = 2000$, $n = 500$, $r = 15$ for population B respectively. Then the bias and M.S.E for suggested estimators under design I and II, using the expressions of bias and M.S.E. derived in Section 5 for suggested estimators are shown in table 7.1 and 7.2 respectively. The bias and M.S.E. for Ahmed's estimators (see Appendix A) are displayed in table 7.3 for population A and B respectively.

Table 7.1 Bias and MSE (Population A)

Estimators	DESIGN I		DESIGN II	
	Bias	MSE	Bias	MSE
$t_4^{'}$	0	10.41747	0	12.31328
$t_5^{'}$	0	36.99100	0	36.78069
$t_6^{'}$	0	10.91418	0	11.29167

Table 7.2 Bias and MSE (Population B)

Estimators	DESIGN I		DESIGN II	
	Bias	MSE	Bias	MSE
$t_4^{'}$	0	16403.58	0	16518.98
$t_5^{'}$	0	22261.45	0	22339.40
$t_6^{'}$	0	16300.30	0	16384.03

Table 7.3 Bias and MSE for Ahmed's Estimators (Population A and B)

Estimators	Population A		Population B	
	Bias	MSE	Bias	MSE
t_4	0	12.73984	0	16531.89
t_5	0	35.83645	0	22319.77
t_6	0	9.759633	0	16358.62

The sampling efficiency of suggested estimators under design I and II over Ahmed's estimators is defined as:

$$E_i = \frac{Opt\left[M\left(t_i^{'}\right)_j\right]}{Opt\left[M\left(t_i\right)\right]}; \qquad i = 4,5,6; \qquad j = I, II \qquad \qquad \ldots(7.1)$$

The efficiency for population A and B respectively given in table 7.4.

Table 7.4 Efficiency of Suggested Estimators in Design *I* and *II* over Ahmed's Estimators

Efficiency	Population A		Population B	
	Design *I*	Design *II*	Design *I*	Design *II*
E_4	0.817709	0.966518	0.992239	0.999219
E_5	1.032217	1.026349	0.997387	1.000879
E_6	1.118298	1.156977	0.996435	1.001553

8. DISCUSSIONS:

The idea of two-phase sampling is used while considering the auxiliary population mean is unknown. Some strategies are suggested for missing observations in Section 4 and the estimators of population mean are derived. Properties of estimators like bias and m.s.e are discussed in the Section 5 and the optimum value of parameters for minimum mean squared error is obtained as well in the same section. Ahmed's estimators are considered for relative comparison. 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 under design *I* and *II* over Ahmed's estimators is obtained and suggested strategy is found very close with Ahmed et al. (2006) when $\overline{X}$ unknown.

9. CONCLUSIONS:

The proposed estimators are useful when some observations are missing in the sample and population mean of auxiliary information is unknown. Obviously from Table 7.1 and 7.2, all suggested estimators are better in design *I* than design *II* i.e. the design *I* is better than design *II*. The table 7.4 shows that the suggested estimators $t_5^{'}$ and $t_6^{'}$ are very close with Ahmed's estimators and may be used to estimate the population mean while population parameter of auxiliary information is unknown.

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. **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.
3. **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.
4. **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 Ame. Stat. Asso. (Social Survey Res. Methods Sec.), 384-389.
5. **Rao, J. N. K. and Sitter, R. R. (1995):** *Variance estimation under two-phase sampling with application to imputation for missing data,* Biometrica, 82, 453-460.
6. **Sande, I. G. (1979):** *A personal view of hot deck approach to automatic edit and imputation.* Journal Imputation Procedures. Survey Methodology, 5, 238-246.
7. **Shukla, D. (2002):** *F-T estimator under two-phase sampling,* Metron, 59, 1-2, 253-263.
8. **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.
9. **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.
10. **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.
11. **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 (IJMER), 1, 2, 580-596.

12. Singh, S. and Horn, S. (2000): *Compromised imputation in survey sampling,* Metrika, 51, 266-276.

APPENDIX – A

Proposed Methods of Ahmed et al. (2006):

Ahmed et al. (2006) proposed some imputation methods and derived their properties. Authors are discussing with three methods of them. Let y_{ji} denotes the ith available observation for the jth imputation and b_i , $i = 1,2,3$ is a suitably chosen constant, such that the variance the resultant estimator is minimum. Imputation methods are :

(1)
$$y_{4i} = \begin{cases} y_i & if \quad i \in R \\ \left[\overline{y}_r + b_1\left(\overline{x}_i - \overline{x}_r\right)\right] & if \quad i \in R^C \end{cases}$$
…(1)

under this strategy, the point estimator of $\overline{Y}$ is $t_4 = \overline{y}_r + b_1\left(\overline{x} - \overline{x}_r\right)$ …(2)

Theorem: The bias, variance and minimum variance at $b_1 = \rho \dfrac{S_{XY}}{S_X^2}$ of t_4 is given by

(i) $B[t_4] = 0$ …(3)

(ii) $V(t_4) = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)S_Y^2 + b_1^2\left(\dfrac{1}{r} - \dfrac{1}{n}\right)S_X^2 - 2b_1\left(\dfrac{1}{r} - \dfrac{1}{n}\right)S_{XY}$ …(4)

(iii) $V(t_4)_{min} = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)S_Y^2 - \left(\dfrac{1}{r} - \dfrac{1}{n}\right)\dfrac{S_{XY}^2}{S_X^2}$ …(5)

(2)
$$y_{5i} = \begin{cases} y_i & if \quad i \in R \\ \overline{y}_r + \dfrac{nb_2}{(n-r)}\left(\overline{X} - \overline{x}\right) & if \quad i \in R^C \end{cases}$$
…(6)

under this strategy, the point estimator of $\overline{Y}$ is $t_5 = \overline{y}_r + b_2\left(\overline{X} - \overline{x}\right)$ …(7)

Theorem: The bias, variance and minimum variance at $b_2 = \rho \dfrac{S_{XY}}{S_X^2}$ of t_5 is given by

(i) $B[t_5] = 0$ …(8)

(ii) $V(t_5) = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)S_Y^2 + b_2^2\left(\dfrac{1}{n} - \dfrac{1}{N}\right)S_X^2 - 2b_2\left(\dfrac{1}{n} - \dfrac{1}{N}\right)S_{XY}$ …(9)

(iii) $V(t_5)_{min} = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)S_Y^2 - \left(\dfrac{1}{n} - \dfrac{1}{N}\right)\dfrac{S_{XY}^2}{S_X^2}$ …(10)

(3)
$$y_{6i} = \begin{cases} y_i & if \quad i \in R \\ \overline{y}_r + \dfrac{nb_3}{(n-r)}\left(\overline{X} - \overline{x}_r\right) & if \quad i \in R^C \end{cases}$$
…(11)

under this , the estimator of $\overline{Y}$ is $t_6 = \overline{y}_r + b_3\left(\overline{X} - \overline{x}_r\right)$ …(12)

Theorem: The bias, variance and minimum variance at $b_3 = \rho \dfrac{S_{XY}}{S_X^2}$ of t_6 is given by

(i) $B[t_6] = 0$ …(13)

(ii) $V(t_6) = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)\left(S_Y^2 + b_3^2 S_X^2 - 2b_3 S_{XY}\right)$ …(14)

(iii) $V(t_6)_{min} = \left(\dfrac{1}{r} - \dfrac{1}{N}\right)S_Y^2\left(1 - \rho^2\right)$ …(15)

Secured Medical Data Publication & Measure the Privacy "Closeness" Using Earth Mover Distance (EMD)

Krishna.V[#], Santhana Lakshmi. S[*]

[#]PG Student, Department of Computer Science
[*] Department of Computer Science
[#*]Coimbatore Institute of Engineering and Technology
Coimbatore, India

Abstract: Privacy requirement for publishing microdata needs equivalent class contains at least k-records. Recent research for releasing microdata are k-anonymity and l-diversity this two method are used to limit only the identity disclosure it is not sufficient for limiting the attribute disclosure. so we are undergoing a new privacy technique known a t-closeness also we are going to measure the privacy that is finding the distance measure between the two probability distribution by using the Earth Mover's Distance a new distance measure technique.

I.INTRODUCTION

Data Publishing is the leading publisher for local communities and business. Data Publishing plays a vital role in hospitals, government agencies, insurance companies and all other business where data would like to release for the purpose of analysis and research purpose, Now a day's society is experiencing exponential growth in the number and variety of data collections containing persons specific information, for example publishing medical data is very much important for analysis and in research areas. Mainly for this purpose the data are stored in table, each table consists of rows and columns, each row consists of records corresponding to one individual and each record has a number of attributes. There are three types of attributes they are 1) **Explicit identifier** are the attribute that clearly identifies the individual, e.g., Patient name, Patient Number. 2) **Quasi identifier** are the attributes whose values that can be taken together can potentially identify an individual, e.g., Zip code, Date-of-birth, Gender. 3) **Sensitive attributes** are the attributes that are consider being more sensitive when releasing a micro data, it is very much necessary to prevent the sensitive information of the individuals from being disclosed, e.g., Disease of an individual is the more sensitive information in medical data publishing.

Privacy Preservation is the main aim of data publishing, however micro data contains more sensitive information it is very necessary for the owner to protect those information i.e.) guaranteeing the privacy of individual by ensuring that their sensitive information is not disclosed. Basically there are two types of disclosures they are as 1) **Identity disclosure** is a type of disclosure when an individual is linked to a particular record in the released table. Once if there is an occurrence of identity disclosure in the released table it is very easy to identify the details of the particular individual. 2) **Attribute disclosure** is another

type of disclosure it occurs when the new information of some individual is revealed i.e., the released data make possible to know the characteristics of a particular individual more accurately. Identity disclosure mainly leads to attribute disclosure but attribute disclosure may occur with or without the occurrence of identity disclosure. It has been recognized that it may cause harm even if there is a disclosure of the false attribute information and also if the perception is incorrect.

Our main aim in data publishing is to limit the disclosure risk of the table that is to be published and this can be achieved by anonymizing the data before release. Anonymization is the technique in which the explicit identifiers are removed, but this is not enough for preserving the published data because the adversary is already having the quasi identifier values in the table. Generalization is the common anonymization approach.

It is necessary to measure the risk of an anonymized table to limit the disclosure risk. For this two of them Samarati and Sweeney introduced a technique known as k-anonymity, property that captures the protection of microdata table with respect to possible re-identification of respondents to which the data refer. It prevents the identity disclosure but it is insufficient to prevent attribute disclosure. To overcome this limitation of k-anonymity , Machanavajjhala introduced a notion of privacy called l-diversity i.e., which it requires that the distribution of sensitive attribute in each equivalence class has at least *l* well represented values. This l-diversity also has some problem that which mainly deals with the limitation of assumption in adversarial knowledge i.e., it is possible for an adversary to gain information about the sensitive attributes as long as he has the information about the global distribution of this attribute.

In this we are going to propose a novel privacy notion called "Closeness" At first we are going to formalize the idea of the base model t-closeness requires that the distribution of sensitive attribute in any equivalence class to be close to the distribution of the attribute in the overall table (i.e., given that the distance between both the distribution should not be more than the threshold t) this can effectively reduce the amount of individual specific information that an observer can learn. We are going to propose an flexible privacy model called (n,t)-closeness. With this we are also going to find the distance between the values of sensitive attribute by using Earth Mover Distance metric(EMD).

II. K-Anonymity

Definition 1 (The k-anonymity principle). *Each release of data must be in such a way that every combination of the values of quasi-identifiers can be indistinctly matched to at least for k respondents.*

If the information for each person contained in the release cannot be distinguished from at least K-1 individuals whose information also appear in the release. It is a type of protection which is provided in a simple and easy way to understand. K-anonymity mainly deals with two factors Suppression and Generalization.

Suppression can replace individual attribute with a * and generalization will replace the individual attributes with a border category for example, consider the age of a person is 35 then in generalization it will be replaced as [30 – 40], while this K-anonymity is sufficient only to protect identity disclosure, but it is not sufficient to protect attribute disclosure. If a table satisfies K-anonymity for some value k, then if anyone knows the quasi identifier value of any particular person then it will be easy to find the details of that individual. Two attacks are identified in this k-anonymity they are Homogeneity attack and Background knowledge attack. K-Anonymity can create groups that leak the information due to the lack of diversity in the sensitive attribute, this may cause both the homogeneity and background knowledge attack.

Table 1 Original Patient Table

S.no	Zip code	Age	Nationality	Disease
1	13053	28	Russian	Heart
2	13068	29	Indian	Cancer
3	13068	21	Japanese	Viral
4	14858	50	American	Viral
5	14853	55	Russian	Cancer
6	14853	47	Indian	Heart
7	14850	33	American	Heart
8	14850	31	Indian	Cancer
9	14853	39	Russian	Heart

Example 1. From the above two table, Table 1 is the original data table and Table 2 is an anonymized version of it satisfying 3-anonymity, in this table disease is consider to be the sensitive attribute. Suppose Alice knows that Bob is 29 year old man living in ZIP 13068 and Bob's record is in table. From the Table 2 Alice can confirm that Bob corresponds to one of the first three records, and must have heart disease, this is called homogeneity attack.

Suppose by knowing Carl's age is 39 in the ZIP code 1305, Alice can conclude that Carl correspond to a record in the last equivalence class in the Table 2. Also Alice knows that Carl has very low risk of heart disease, this is background knowledge attack i.e., this enables Alice to conclude that Carl most likely has cancer.

Table 2 A 3-Anonymous version of Table

S.no	Zip code	Age	Nationality	Disease
1	130**	2*	*	Heart
2	130**	2*	*	Cancer
3	130**	2*	*	Viral
4	148**	[45-60]	*	Viral
5	148**	[45-60]	*	Cancer
6	148**	[45-60]	*	Heart
7	1485*	3*	*	Heart
8	1485*	3*	*	Cancer
9	1485*	3*	*	Heart

l-diversity

To address the limitation of K-Anonymity recently introduced a new notion of privacy known as *l*-diversity. which requires that the distribution of a sensitive attribute in each equivalence class has at least *l* "well represented" values.

Principle of *l*-diversity

A q-block is *l*-diverse if it contains at least *l* "well represented" values for the sensitive attribute S. A table is *l*-diverse if every q-block is *l*-diverse.

l-diversity Instantiations

l-diversity consists of certain instantiations and are stated by Machanavajjhala.

1. **Distinct *l*-diversity.** Each equivalence class has at least *l* "well represented" sensitive values. The drawback of this distinct *l*-diversity is it does not prevent probabilistic interface attacks. For example in an equivalent class. The table contains ten tuples in the sensitive area disease in that consider that one of them is "flu" and one is "Cancer" and the rest eight are "Heart Disease" this satisfies 3-diversity rule, but the attacker can still confirm that the target person's disease is "Heart Disease" with the accuracy of 80%.

2. **Entropy *l*-diversity.** Each equivalence class not only must have enough different sensitive values, but also the different sensitive values must be distributed evenly, it means the entropy of the distribution of sensitive values in each class is at least log(l). the disadvantage of this is sometimes it may be too restrictive. i.e., in a table some values are more common means then the entire table entropy is too low. This may cause less conservative notion of *l*-diversity.

3. **Recursive (c,*l*)-diversity.** This recursive (c,*l*)-diversity can be interpreted in terms of adversarial background knowledge. This mainly protect against all adversaries who posses almost *l*-2 diversity. The main drawback of this diversity is that the most frequent value won't appear more frequently, also the less frequent value does not appear too rarely. If $r1 < c(rl + rl+1 + ... + rm)$ then the table is said to have recursive (c,*l*)-diversity if all of its equivalence classes have recursive (c,*l*)-diversity.

Limitations of *l*-diversity

We are using *l*-diversity in order to overcome the disadvantage of k-anonymity beyond protecting against the attribute disclosure. the main disadvantage of *l*-diversity is that it won't consider the overall distribution of the sensitive values, *l*-diversity is difficult to achieve and also it does not provide sufficient protection against attribute disclosure.

Example 2. Suppose there are 10000 records in total in that if 99 percent is negative and only 1 percent is positive means then the two values have very difficult degrees of sensitivity. i.e., in this one won't mind for being known to test for negative, because one is same as 99 percent of the population, but one would not want to test for positive. In this case 2-diversity does not provide sufficient privacy protection.

This *l*-diversity is insufficient to provide attribute disclosure has two types of attack namely skewness attack and similarity attack.

Skewness attack.

Consider an example that one equivalent class has equal number of positive and negative records means it will satisfy distinct 2-diversity, entropy 2-diversity and (c,2)-diversity, by this we can consider that the 50 percent of the possibility be positive and the other 50 percent be negative.

Similarity attack

Sensitive attribute value in an equivalent class are said to be distinct but also semantically similar, in this case an adversary can learn the important information from the table and is said to be as similarity attack and is shown in table 3 and 4.

Table 3 Original Salary/Disease Table

	Zip code	Age	Salary	Disease
1	47671	25	4k	Gastric
2	47603	23	3k	Gastric
3	47677	24	5k	Cancer
4	47905	45	11k	Gastric
5	47980	53	6k	Bronchitis
6	47906	42	8k	Bronchitis
7	47603	32	10k	Flu
8	47609	39	9k	Pneumonia
9	47607	35	10k	Cancer

III. NEW PRIVACY MEASURE: t-closeness

t-closeness is said to be as a new privacy measure which is said to be as the distribution of a sensitive attribute in any equivalence class to the distribution of sensitive attribute in the overall table. It is an enhancement model of *l*-diversity. Generally privacy can be measured by the information gained by the observer and this can be measured or calculated by subtracting the posterior belief and the prior belif of the observer.

Table 4 A 3-Diverse Version of Table

	Zipcode	Age	Salary	Disease
1	476**	2*	4k	Gastric
2	476**	2*	3k	Gastric
3	476**	2*	5k	Cancer
4	479**	[40-50]	11k	Gastric
5	479**	[40-50]	6k	Bronchitis
6	479**	[40-50]	8k	Bronchitis
7	476**	3*	10k	Flu
8	476**	3*	9k	Pneumonia
9	476**	3*	10k	Cancer

The *l*-diversity requirement is motivated by limiting the difference between the posterior and prior belief. In this we are not going to limit the information gained by the observer about the whole population but we are going to limit the extent to which an observer can learn additional information about the specific individual. *An equivalence class is said to have t-closeness if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole table should not be no more than threshold t.* Based on the analysis, we propose a more flexible ***privacy model called (n, t) - closeness***, which requires that the distribution in any equivalence class is close to the distribution in a **large-enough equivalence class** (contains at least n records) with respect to the sensitive attribute.

t-closeness can be performed based on the utility analysis by using the Greedy algorithm, the primary technique used for the generation of k-anonymous and l-diversity table from the original data set table are to generalize the quasi-identifier values and the sensitive values that are got from former and latter. This generalization can be performed in two ways namely they are of for the semantic data and the numeric data. In semantic data the data moves up by generalization hierarchy either by implied or by supplied. In numeric data a specific case of an implied generalization hierarchy is used.

Privacy can be measured by the information gain by an observer, i.e., at first the observer may have the prior belief about the sensitive attribute value before seeing the released table, and then after seeing the released table the observer may get the posterior belief about the sensitive attribute value. Prior belief is the distribution of sensitive attribute value in the whole table and it is represented by Q. Posterior belief is distribution of sensitive attribute value in a single class and this can be represented by P. From this we know that the information gain can be represented as the difference got from the posterior belief to the prior belief.

Information Gain = Posterior belief – Prior belief (or)
Information Gain = D[P,Q]

Architecture
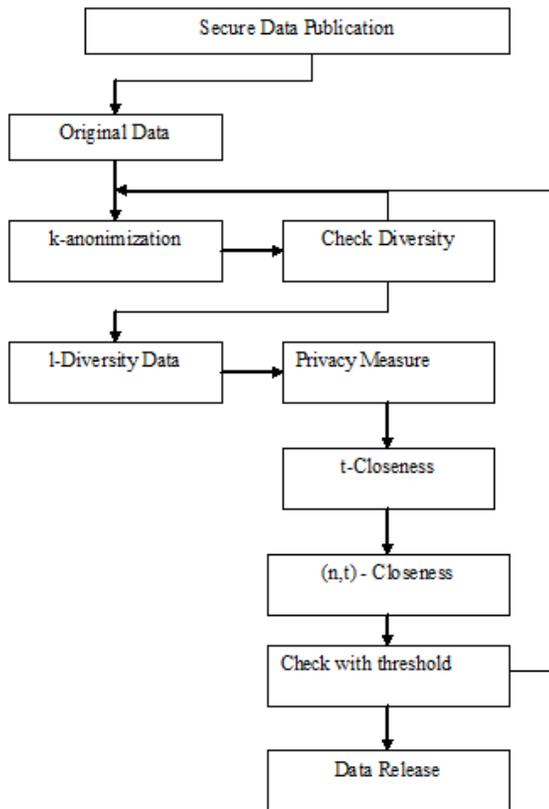
Figure 1Architecture of Microdata Release

t-Closeness Principle algorithm
input: P and Q is partitioned into r partitions as $\{P_1, P_2,\ldots\ p_r\}$ and $\{Q_1,Q_2,\ldots Q_r\}$, EC is Each Class, t is the threshold value.

Output: true if (n,t)-closeness is satisfied, otherwise false.

Information Gain = D[P,Q]

An EC is t-Closeness if $D[P,Q] \leq t$

An Table is said to be as t-Closeness if and only if all the EC has t-Closeness

If D[P,Q] is ↓, then the information gained by the observer will ↓ privacy risk will also get ↓

If D[P,Q] ↑, then the information gained by the observer will also ↑ the benefit of the published data

---

Where in this P → Posterior Belief

Q →Prior Belief and

D →Difference

IV.DISTANCE MEASURES

Distance Measure is the technique which is used for measuring the distance that is the security, which is must satisfy some five properties such as identifying of indiscernible, Non negativity, Probability Scaling, Zero Probability definability and semantic awareness. Identity of indiscernible is that no information is gained if an adversary does not change its belief it is generally represented as   D[P,Q] = 0. Non negativity means if an adversary is gaining non negative information then it can be represented as $D[P,Q] \geq 0$. Probability Scaling means D[P,Q] should compulsorily reflect the difference. Zero probability distribution is the well defined zero probability values in P and Q and semantic awareness is that if the values of P and Q are of having semantic meaning means then D[P,Q] must reflect the semantic difference among that two different values.

In general we are planning to measure the distance with the help of two types of algorithms namely they are of two ways they are Similarity Measure and Earth Mover's Distance which is mainly used to calculate or to measure the privacy.

Similarity Measure:

Similarity Measure is the technique that which mainly allows similar evaluation of the encrypted policies. This technique will relies the existing encryption method that which mainly allows for the numerical data that are present in the table. Similarity Measure is the function that maps a pair of attributes to the interval [0,1]. It captures the intuitive notion of two values being "similar." Generally similar attributes will behave like an indicator function.

DE-ANONYMIZATION ALGORITHM:

De-anonymization algorithm uses matching function and scoring function. Scoring function assigns the numerical value of the data table and matching function mainly deals with the algorithm applied by the adversary to determine the scores by using the set of matches. Finally record selection selects one "best guess" record.

Earth Mover's Distance:

The EMD computes the distance between two distributions, which are represented by signatures. The signatures are sets of weighted features that capture the distributions. The features can be of any type and in any number of dimensions, and are defined by the user. The EMD is defined as the minimum amount of work needed to change one signature into the other. The notion of "work" is based on the user-defined ground distance which is the distance between two features. The size of the two signatures can be different. Also, the sum of weights of one signature can be different than the sum of weights of the other (partial match). Because of this, the EMD is normalized by the smaller sum.

V. CONCLUSION

The method we used here will surely reduce the disclosure risk and also they provide the high level security which is very much useful in microdata publishing. t-closeness is the more flexible privacy model that which provide or achieves the better balance between the privacy and utility. t-closeness removing an outlier may smooth a distribution and it bring it much closer to the overall distribution. For measuring the privacy we use similarity measure and the Earth Mover's Distance for performing all this process we use generalization and suppression techniques.

REFERENCES

[1]   C. Aggarwal, "On k-Anonymity and the Curse   of Dimensionality," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 901-909,2005.
[2]   G. Aggarwal, T. Feder, K. Kenthapadi, S. Khuller, R. Panigrahy, D.Thomas, and A. Zhu, "Achieving Anonymity via Clustering,"Proc. ACM Symp. Principles of Database Systems (PODS), pp. 153-162, 2006.
[3]   R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network Flows: Theory, Algorithms, and Applications. Prentice-Hall, Inc., 1993.
[4]   R.J. Bayardo and R. Agrawal, "Data Privacy through Optimal k-Anonymization," Proc. Int'l Conf. Data Eng. (ICDE), pp. 217-228,2005.
[5]   F. Bacchus, A. Grove, J.Y. Halpern, and D.   Koller, "From Statisticsto Beliefs," Proc. Nat'l Conf. Artificial Intelligence (AAAI), pp. 602-608, 1992.
[6]   J.-W. Byun, Y. Sohn, E. Bertino, and N. Li, "Secure Anonymization for Incremental Datasets," Proc. VLDB Workshop Secure Data Management (SDM), pp. 48-63, 2006.
[7]   B.-C. Chen, K. LeFevre, and R. Ramakrishnan,   "Privacy Skyline: Privacy with Multidimensional Adversarial Knowledge," Proc. Int'l Conf. Very Large Data Bases   (VLDB), pp. 770-781, 2007.
[8]   G.T. Duncan and D. Lambert, "Disclosure-   Limited Data Dissemination,"J. Am. Statistical Assoc., vol. 81, pp. 10-28, 1986.
[9]   D. Kifer and J. Gehrke, "Injecting Utility into Anonymized Datasets," Proc. ACM SIGMOD, pp. 217-228, 2006.
[10]   D. Lambert, "Measures of Disclosure Risk and Harm," J. Official Statistics, vol. 9, pp. 313-331, 1993.
[11]   K. LeFevre, D. DeWitt, and R. Ramakrishnan, "Workload-Aware Anonymization," Proc. ACM SIGKDD, pp. 277-286, 2006.
[12]   T. Li and N. Li, "Injector: Mining Background Knowledge for Data Anonymization," Proc. Int'l Conf. Data Eng. (ICDE), 2008.
[21] T. Li and N. Li, "Towards Optimal k- Anonymization," Data and Knowledge Eng., vol. 65, pp. 22-39, 2008.
[22] T.M. Truta and B. Vinay, "Privacy Protection: P-Sensitive k-Anonymity Property," Proc. Int'l Workshop Privacy Data Management (ICDE Workshops), 2006.
[23]   A. Asuncion and D.J. Newman, UCI Machine Learning Repository, 2007.

ABOUT THE AUTHORS

The Author, Krishna.V is a final year student doing Master of Engineering in Computer Science at Coimbatore Institute of Engineering and Technology and has a Bachelor of Engineering degree in Computer Science and Engineering   from Ponjesly College of Engineering, Nagercoil. My Research area is Data Mining.

Ms. S. Santhana Lakshmi received her M.E., degree in Computer Science and Engineering from VLB Janakiammal college of Engineering .She is currently working as assistant professor in Coimbatore Institute of Engineering and Technology. Her Research area is Network Security and Data Mining.

DESIGN AND ANALYSIS OF
COMPOSITE LEAF SPRING IN LIGHT VEHICLE

M.VENKATESAN [1], D.HELMEN DEVARAJ [2],

* Assistant Professor, Department of Mechanical Engineering, Sona College of Technology, Salem-5, India.
**M.E – Engineering Design, Department of Mechanical Engineering, Sona College of Technology, Salem-5, India.

ABSTRACT:
This project describes design and experimental analysis of composite leaf spring made of glass fiber reinforced polymer. The objective is to compare the load carrying capacity, stiffness and weight savings of composite leaf spring with that of steel leaf spring. The design constraints are stresses and deflections. The dimensions of an existing conventional steel leaf spring of a light commercial vehicle are taken. Same dimensions of conventional leaf spring are used to fabricate a composite multi leaf spring using E-Glass/Epoxy unidirectional laminates. Static analysis of 2-D model of conventional leaf spring is also performed using ANSYS 10 and compared with experimental results. Finite element analysis with full load on 3-D model of composite multi leaf spring is done using ANSYS 10 and the analytical results are compared with experimental results. Compared to steel spring, the composite leaf spring is found to have 67.35% lesser stress, 64.95% higher stiffness and 126.98% higher natural frequency than that of existing steel leaf spring. A weight reduction of 76.4% is achieved by using optimized composite leaf spring.

Keywords – **Composite materials, design constrains, leaf spring, material property, and static analysis.**

1. INTRODUCTION:

In order to conserve natural resources and economize energy, weight reduction has been the main focus of automobile manufacturers in the present scenario. Weight reduction can be achieved primarily by the introduction of better material, design optimization and better manufacturing processes. The suspension leaf spring is one of the potential items for weight reduction in automobiles as it accounts for 10% - 20% of the unstrung weight. This achieves the vehicle with more fuel efficiency and improved riding qualities. The introduction of composite materials was made it possible to reduce the weight of leaf spring without any reduction on load carrying capacity and stiffness.

Since, the composite materials have more elastic strain energy storage capacity and high strength to weight ratio as compared with those of steel, multi-leaf steel springs are being replaced by mono-leaf composite springs. The composite material offer opportunities for substantial weight saving but not always are cost-effective over their steel counter parts.

The leaf spring should absorb the vertical vibrations and impacts due to road irregularities by means of variations in the spring deflection so that the potential Energy is stored in spring as strain energy and then released slowly. So, increasing the energy storage capability of a leaf spring ensures a more compliant suspension system. According to the studies made a material with maximum strength and minimum modulus of elasticity in the longitudinal direction is the most suitable material for a leaf spring. Fortunately, composites have these characteristics.

Fatigue failure is the predominant mode of in-service failure of many automobile components. This is due to the fact that the automobile components are subjected to variety of fatigue loads like shocks caused due to road irregularities traced by the road wheels, the sudden loads due to the wheel traveling over the bumps etc. The leaf springs are more affected due to fatigue loads, as they are apart of the unstrung mass of the automobile.

The fatigue behavior of Glass Fiber Reinforced Plastic (GFRP) epoxy composite materials has been studied in the past. Theoretical equation for predicting fatigue life is formulated using fatigue modulus and its degrading rate. This relation is simplified by strain failure criterion for practical application. A prediction method for the fatigue strength of composite structures at an arbitrary combination of frequency, stress ratio and temperature has been presented. These studies are limited to mono-leaf springs only.

In the present work, a seven-leaf steel spring used in passenger cars is replaced with a composite multi leaf spring made of glass/epoxy composites. The dimensions and the number of leaves for both steel leaf spring and composite leaf springs are considered to be the same. The primary objective is to compare their load carrying capacity, stiffness and weight savings of composite leaf spring. Finally, fatigue life of steel and composite leaf spring is also predicted using life data.

2. LITERATURE REVIEW:

2.1 Ballinger C.A. – Getting Composites into Construction, Reinforced Plastics, 1995.

Composite leaf spring in the early 60 failed to yield the production facility because of inconsistent fatigue

performance and absence of strong need for mass reduction. Researches in the area of automobile components have been receiving considerable attention now. Particularly the automobile manufacturers and parts makers have been attempting to reduce the weight of the vehicles in recent years. Emphasis of vehicles weight reduction in 1978 justified taking a new look at composite springs. Studies are made to demonstrate viability and potential of FRP in automotive structural application.

The development of a lit flex suspension leaf spring is first achieved. Based on consideration of chipping resistance base part resistance and fatigue resistance, a carbon glass fiber hybrid laminated spring is constructed. A general discussion on analysis and design of constant width, variable thickness, and composite leaf spring is presented. The fundamental characteristics of the double tapered FRP beam are evaluated for leaf spring application.

Recent developments have been achieved in the field of materials improvement and quality assured for composite leaf springs based on microstructure mechanism. All these literature report that the cost of composite; leaf spring is higher than that of steel leaf spring. Hence an attempt has been made to fabricate the composite leaf spring with the same cost as that of steel leaf spring.

2.2 Miravete.A, Castejon. L, Bielsa.J, Bernal.E - Analysis and Prediction of large composite Structures, 1990.

Material properties and design of composite structures are reported in many literatures. Very little information is available in connection with finite element analysis of leaf spring in the literature, than too in 2D analysis of leaf spring. At the same time, the literature available regarding experimental stress analysis more.

The experimental procedures are described in national and international standards. Recent emphasis on mass reduction and developments in materials synthesis and processing technology has led to proven production worthy vehicle equipment..

3. SPECIFICATION OF THE PROBLEM:

The objective of the present work is to design, analyses, Glass Fiber/Epoxy complete composite leaf spring with out end joints and composite leaf spring using bonded end joints using hand-lay up technique. This is an alternative, efficient and economical method over wet filament-winding technique.

4. LEAF SPRINGS:

Leaf springs also known as flat spring are made out of flat plates. Leaf springs are designed two ways: multi-leaf and mono-leaf. The leaf springs may carry loads, brake torque, driving torque, etc... In addition to shocks.

The multi-leaf spring is made of several steel plates of different lengths stacked together. During normal operation, the spring compresses to absorb road shock. The leaf springs bend and slide on each other allowing suspension movement.

4.1 Construction of Leaf Spring:

The leaves are usually given an initial curvature or cambered so that they will tend to straighten under the load. The leaves are held together by means of band shrunk around them at the centre or by a bolt passing through center. Since, the band exerts stiffening and strengthening effect, therefore effective length of the spring for bending will be overall length of the spring minus width of the band.

In case of a center bolt, two-third distance between centers of U-bolt should be subtracted from the overall length of the spring in order to find effective length. The spring is clamped to the axle housing by means of U-bolts. The longest leaf known as main leaf or master leaf has its ends formed in the Shape of an eye through which the bolts are passed to secure the spring to its supports.

The other leaves of the spring are known as graduated leaves. In order to prevent digging in the adjacent leaves, the ends of the graduated leaves are trimmed in various forms. Rebound clips are located at intermediate positions in the length of the spring, so that the graduated leaves also share the stress induced in the full length leaves when the spring rebounds.
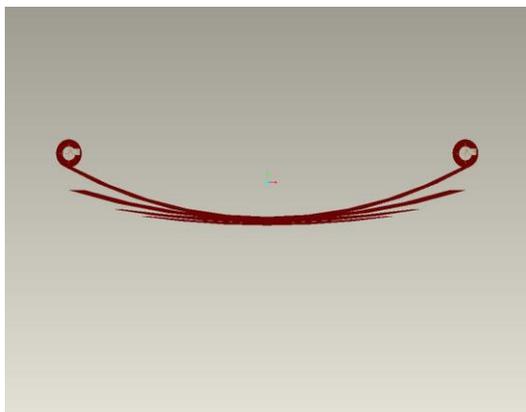
Fig. No: 1 – Front View.
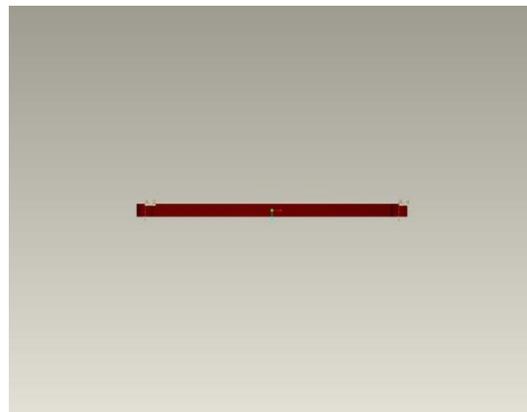
Fig. No: 2 – Top View.

4.2 Materials for Leaf Springs:

The material used for leaf springs is usually a plain carbon steel having 0.90 to 1.0% carbon. The leaves are heat treated after the forming process. The heat treatment of spring steel products greater strength and therefore greater load capacity, greater range of deflection and better fatigue properties.

5. COMPOSITE MATERIAL:
5.1 Characteristics:

A composite material is defined as a material composed of two or more constituents combined on a macroscopic scale by mechanical and chemical bonds.

Typical composite materials are composed of inclusions suspended in a matrix. The constituents retain their identities in the composite. Normally the components can be physically identified and there is an interface between them.

Many composite materials offer a combination of strength and modulus that are either comparable to or better than any traditional metallic materials. Because of their low specific gravities, the strength weight-ratio and modulus weight-ratios of these composite materials are markedly superior to those of metallic materials.

The fatigue strength weight ratios as well as fatigue damage tolerances of many composite laminates excellent. For these reasons, fiber composite have emerged as a major class of structural material and are either used or being considered as substitutions for metal in many weight-critical components in aerospace, automotive and other industries.

Another unique characteristic of many fiber reinforced composites is their high internal damping. This leads to better vibration energy absorption within the material and results in reduced transmission of noise land vibration to neighboring structures.

High damping capacity of composite materials can be beneficial in many automotive applications in which noise, vibration, and hardness is a critical issue for passenger comfort.

Among the other environmental factors that may cause degradation in the mechanical properties of some polymeric matrix composites are elevated temperatures, corrosive fluids, and ultraviolet rays.

In many metal matrix composites, oxidation of the matrix well as adverse chemical reaction between fibers and matrix are of great concern at high temperature applications.

For automobiles*:* 50Cr 1, 50 Cr 1 V 23, and 55 Si 2 Mn 90 all used in hardened and tempered state.
For rail road springs: C 55 (water - hardened),
C 75 (oil-hardened), 40 Si 2 Mn 90 (water-hardened) and 55 Si 2 M N90 (oil-hardened).

5.2 Applications:

Commercial and industrial applications of composite s are so varied that it is impossible to list them all. The major structural application areas, which include aircraft, space, automotive, sporting goods, and marine engineering. A potential for weight saving with composites exists in many engineering field. The first major structural application of composite is the corvette rear leaf spring in 1981. A uni-leaf E-glass – reinforced epoxy has been used to replace a ten-leaf steel spring with nearly an 80 % weight savings.

Other structural chassis components, such as drive shafts and road wheels, have been successfully tested in the laboratories and are currently being developed for future cars and vans.

The metal matrix composites containing either continuous or discontinuous fiber reinforcements, the latter being in the form of whiskers that are approximately 0.1-0.5 μm in diameter and have a length to diameter ratio up to 200.

Particulate-reinforced metal matrix composites containing either particles or platelet that ranges in size from 0.5 to 100 μm. Dispersion-strengthened metal matrix composites containing particles that are less than 0.1 μm in diameter. And metal matrix composites are such as directionally solidified eutectic alloys.

5.3 Benefits:
i. Weight reduction,
ii. High strength,
iii. Corrosiveness,
iv. Low specific gravity.

6. DESIGN SELECTION:

The leaf spring behaves like a simply supported beam and the flexural analysis is done considering it as a simply supported beam. The simply supported beam is subjected to both bending stress and transverse shear stress. Flexural rigidity is an important parameter in the leaf spring design and test out to increase from two ends to the center.

6.1 Constant Thickness, Varying Width Design:
In this design the thickness is kept constant over the entire length of the leaf spring while the width varies from a minimum at the two ends to a maximum at the center.

6.2 Constant Width, Varying Thickness Design:
In this design the width is kept constant over the entire length of the leaf spring while the thickness varies from a minimum at the two ends to a maximum at the center.

6.3 Constant Cross-Selection Design:

In this design both thickness and width are varied through out the leaf spring such that the cross-section area remains constant along the length of the leaf spring. Out of the above mentioned design concepts. The constant cross-section design method is selected due to the following reasons:Due to its capability for mass production and accommodation of continuous reinforcement of fibers.

Since the cross-section area is constant through out the leaf spring, same quantity of reinforcement fiber and resin can be fed continuously during manufacture. Also this is quite suitable for filament winding process.

7. SPECIFICATION OF EXISTING LEAF SPRING:

Specifications		
1	Total Length of the spring (Eye to Eye)	1540 mm
2	Free Camber (At no load condition)	136mm
3	No. of full length leave (Master Leaf)	01
4	Thickness of leaf	13mm
5	Width of leaf spring	70mm
6	Maximum Load given on spring	3850N
7	Young's Modulus of the spring	22426.09 N/mm^2
8	Weight of the leaf spring	23 kg

8. MATERIAL PROPERTIES OF E-GLASS/EPOXY:

Sl.No	Properties	Value
1	Tensile modulus along X-direction (Ex), MPa	34000
2	Tensile modulus along Y-direction (Ey), MPa	6530
3	Tensile modulus along Z-direction (Ez), MPa	6530
4	Tensile strength of the material, Mpa	900
5	Compressive strength of the material, Mpa	450
6	Shear modulus along XY-direction (Gxy), Mpa	2433
7	Shear modulus along YZ-direction (Gyz), Mpa	1698
8	Shear modulus along ZX-direction (Gzx), Mpa	2433
9	Poisson ratio along XY-direction (Nuxy)	0.217
10	Poisson ratio along YZ-direction (NUyz)	0.366

11	Poisson ratio along ZX-direction (NUzx)	0.217
12	Mass density of the material (⟩), kg/mm3	2.6×10^{-6}
13	Flexural modulus of the material, MPa	40000
14	Flexural strength of the material, MPa	1200

9. THREE-DIMENSIONAL FINITE ELEMENT ANALYSIS:

To design composite leaf spring, a stress analysis was performed using the finite element method done using ANSYS software. Modeling was done for every leaf with eight-node 3D brick element (solid 45) and five-node 3Dcontact element (contact 49) used to represent contact and sliding between adjacent surfaces of leaves. Also, analysis carried out for composite leaf spring with bonded end joints for Glass/Epoxy. The maximum and shear stresses along the adhesive layer were measured; represent FEA results for composite leaf spring (Glass/Epoxy). The maximum and shear stresses along the bonded adhesive layer for glass/epoxy were measured and plotted as shown in Figs3 & 4.
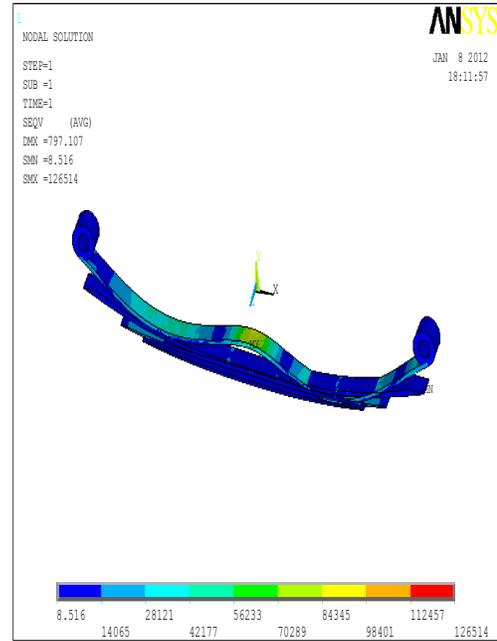
Fig. No: 4
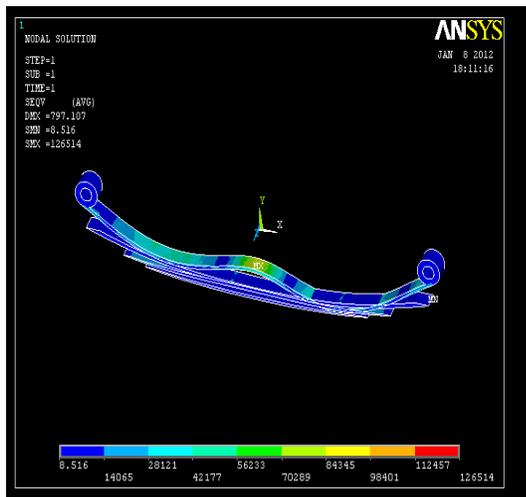
Fig. No: 3
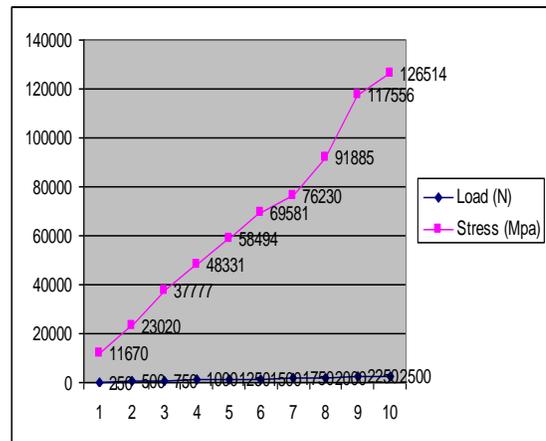
Fig. No: 5 - Stress distribution in adhesive for Glass/Epoxy Leaf spring.
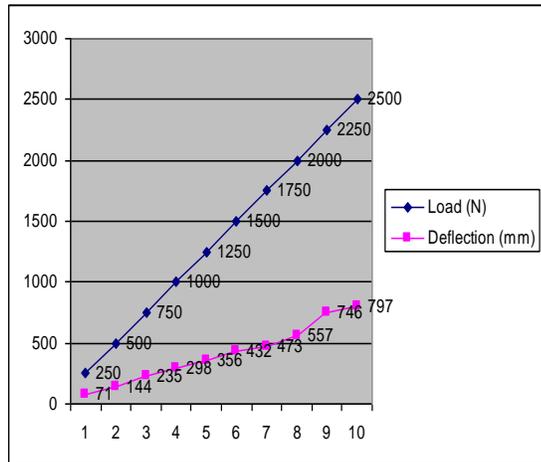
Fig. No: 6 – Deflection distribution in adhesive for Glass/Epoxy Leaf spring.

10. CONCLUSION:

The development of a composite leaf spring having constant cross sectional area, where the stress level at any station in the leaf spring is considered constant due to the parabolic type of the thickness of the spring, has proved to be very effective.

The study demonstrated that composites can be used for leaf springs for light weight vehicles and meet the requirements, together with substantial weight savings.

The 3-D modeling of composite leaf spring is done and analyzed using ANSYS.

A comparative study has been made between composite and steel leaf spring with respect to weight, cost and strength.

From the results, it is observed that the composite leaf spring is lighter and more economical than the conventional steel spring with similar design specifications.

Composite leaf spring reduces the weight by 85 % for E-Glass/Epoxy, over conventional leaf spring.

11. REFERENCES:

11.1 Journal Papers:

[1] Hawang, W., Han, K. S. Fatigue of Composites – Fatigue Modulus Concept and Life Prediction *Journal of Composite Materials, 1986.*

[2] Dharam, C. K. Composite Materials Design and Processes for Automotive Applications. *The ASME Winter Annual Meeting, San Francisco, 1978.*

[3] Springer, George S., Kollar, Laszloa P. Mechanics of Composite Structures. *Cambridge University Press, New York, 2003.*

[4] AL-Qureshi, H. A. Automobile leaf springs from composite materials, *Journal of* Processing *Technol., 2001.*

11.2 Books:

[5] P. Beardmore, Composite *structure for automobiles,* 1986.

[6] R.S. Khurmi, J.K. Kupta. A *text book of Machine Design,* 2000.

[7] R. M. Jones, *Mechanics of Composite Materials.* 2e, McGraw-Hill Book Company, 1990.

[8] K. Tanabe, T. Seino, Y. Kajio, *Characteristics of Carbon/Glass Fiber Reinforced Plastic Leaf Spring*, 1982.

11.3 Chapters in Books:

[9] R.S. Khurmi, J.K. Kupta. A *text book of Machine Design,* 2000, Chapter. 23, Page. No (866-874).

11.4 Proceedings Papers:

[10] Daugherty, R. L. Composite Leaf Springs in Heavy Truck Applications. K. Kawata, T.Akasaka (Eds). *Composite Materials Proceedings of Japan-US Conference Tokyo,1981.*

A MULTI-HOP WIRELESS MBS USING AN ADAPTIVE NETWORK CODE RETRANSMISSION

Yuvaraj.N*, Kalaiselvi.R[#]

*Final ME (CSE), Sri Shakthi Institute Of Engineering and Technology,Anna University,Coimbatore
[#]Asst.Profesor(CSE), Sri Shakthi Institute Of Engineering and Technology,Coimbatore

Abstract:

Network coding has recently attracted attention as a substantial improvement to packet retransmission schemes in wireless multicast broadcast services (MBS). In this paper, it is concentrating on increasing the efficiency of bandwidth and reducing the retransmission of packets on wireless multicast broadcast services. In the existing system there are three main schemes are used. ONCR (Opportunistic network coding retransmission), FNCR (Full network coding retransmission), ANCR (Adaptive network coding retransmission).First two scheme are not efficient because, in ONCR it will transmit the packets which are not intent to the receiver. In FNCR, it decreases the bandwidth efficiency because of retransmitting all the packets in addition with the required packets. In order to overcome the above drawbacks, Adaptive network coded schemes are used to improve the bandwidth efficiency in MBS using a combination of opportunistic and full network coded retransmission using this transmission scheme. It can able to apply it for the single-hop wireless broadcast services. In the proposing system, are applying this transmission scheme for the multi-hop wireless broadcast service. It has the advantage of more effective bandwidth usage and reducing the retransmission.

Index terms- Chromatic number of random graph, Graph coloring, Multicast Broadcast Services(MBS), Full and Opportunistic Network Coded Packet Retransmision.

INTRODUCTION

Multicast Broadcast Services (MBS) have become essential applications that are greatly considered in the design of all future wireless networks due to the increasing demand of an application that are requested by all receivers located in the coverage area of a wireless access node. In multicast, the receivers are interested in receiving only a subset of the packets transmitted by the access node. In broadcast, the receivers are interested in receiving all the packets transmitted by the access node. Due to the high demand on MBS applications and their high bandwidth requirements, it is very important to develop new techniques that can improve the system bandwidth efficiency in order to satisfy these demands with the quality of service.

To achieve a reliable multicast/broadcast, all the receivers must correctly detect all the information packets they requested from the access node. Since wireless communication channels are lossy in general, the guarantee of packet delivery is achieved through packet retransmission using Automatic Repeat Request (ARQ) or Forward Error Correction (FEC). However, both schemes retransmit lost packets separately, which considerably reduces the number of receivers benefiting from each retransmission. This results in more retransmissions and thus low system bandwidth efficiency.

Since these are not efficiency, bandwidth efficiency improvements achieved by network coding [6], several works aimed to exploit it for packet retransmission, as a substitute to ARQ/HARQ in wireless networks. So in the existing system, they introduced two schemes which focus on improving bandwidth and reducing loss of retransmission. Opportunistic Network Coded Retransmission (ONCR) they combine lost packets of different receivers such that some of them recover one of their missing packets upon correct delivery of this combined packet. Second scheme, Full Network Coded Retransmission (FNCR) was proposed to improve wireless multimedia broadcast.

It has been shown that both ONCR and FNCR schemes achieve a considerable gain in bandwidth efficiency compared to ARQ. Each of these two schemes usually outperforms the other in different receiver, demand and feedback settings [1]. The continuous and rapid change of these settings in wireless networks limits the bandwidth efficiency gains if only one scheme is always employed. To select the scheme for the higher bandwidth efficiency, it uses the proposing scheme.

So a new scheme Adaptive Network Coded Retransmission (ANCR) which is a combination of Opportunistic and Full Network Coded Retransmissions. The proposed scheme adaptively selects, between these two schemes, the one that is expected to achieve the better bandwidth efficiency performance. The core contribution in this Adaptive Selection scheme is focusing on derivation of an ONCR performance metric that achieves efficient selection when compared to an appropriate full network coding. This metric is derived by modeling the ONCR graph representation as a random graph and computing its chromatic number using a famous result from random graph theory [7].

The rest of this paper is organized as follows. In Section II, some related works to our problem are summarized. The single-hop wireless MBS system model and its parameters are illustrated in Section III. In Section IV, it has briefly illustrate the ONCR and FNCR schemes in the

general MBS case. Then this paper's main contribution is introduced in Section V by introducing the theoretical foundation and the detailed description of our proposed adaptive scheme.

AN OVERVIEW OF RELATED WORK

A. Network Coding

Since its first introduction in [6], network coding has been a great attraction to several studies as a routing and scheduling scheme that attains maximum information flow in a network. The core of network coding is the idea of packet mixing using several techniques such as packet XOR and linear coding. Both trends have been proposed for a wide range of applications.

B. Index Coding

The index coding includes a sender, a set of receivers, a set of packets and lossless channels between the sender and these receivers. The objective of the index coding problem is to define the packet coding schedule that delivers the requested subsets of packets by each of the receivers with the minimum number of transmissions.

In [2] it has been shown that finding the optimal solution of the index coding problem is NP-hard. Most of these heuristics are different simplifications of suboptimal graph-coloring solution of the index coding problem.

C. Network Coded Retransmission

In [3] and [4], the diversity of received and lost packets at different receivers is broken by using the ONCR scheme as a substitute of ARQ/HARQ, respectively. In [5], a hybrid ARQ-FNCR scheme was proposed for wireless multimedia broadcast. In the concept of network coded retransmissions, minimizing the average packet detection delay and the average sender queue size, respectively, in wireless broadcast.

SYSTEM MODEL AND PARAMETERS

The model consists of a wireless access node, such as a base station which is responsible for delivering multicast or broadcast packets $R=\{R_1, \ldots . R_M\}$ of M receivers. The access node initially transmits a MBS frame consisting $p=\{p_1, \ldots p_N\}$ of N packets. During this phase, each receiver listens to the packets it requested and all correctly received packets are stored in its memory. For each lost packet, each receiver sends a NAK packet to the access node. The access node keeps a table of received and lost packets by all receivers that will refer to as the feedback table. At the end of the initial transmission phase, three sets of packets can be associated with each receiver R_i.

- The Has set (denoted by $\mathbf{H_i}$) is defined as the set of packets correctly received by R_i. This set includes both desired and undesired packets by this receiver.

- The Complementary set (denoted by $\mathbf{C_i}$) is defined as the set of packets that were not correctly received by R_i whether requested or not.

- The Wants set (denoted by $\mathbf{W_i}$) is defined as the set of packets that are both requested and lost by R_i in the initial transmission phase of the current MBS frame.

At the end of the initial transmission phase, a packet retransmission scheme is employed to deliver the lost packets to the receivers that requested them. Afterwards, the whole procedure is re-executed for a new MBS frame.

ONCR AND FNCR SCHEMES

A. ONCR Scheme

The ONCR scheme combines lost packets of different receivers such that some of them recover one of their missing packets upon correct delivery of this combined packet. Each packet combination is performed so as to maximize the number of receivers that directly recover one of their requested and lost packets upon correct reception of this coded packet.

The opportunistic packet coding sequence to minimize the number of retransmissions is equivalent to solving the corresponding index coding problem. Since solving index coding problems is NP-hard, the graph-coloring approximation, proposed in [8], can be used to efficiently implement the ONCR scheme in case of lossless retransmissions.

The graph-coloring implementation of the ONCR scheme starts by generating a graph $\mathbf{G(V,E)}$, in which each packet j € W_i for every I induces a vertex $V_{i,j}$ in the graph.Two vertices $V_{i,j}$ and $V_{k,l}$ in G are connected if one of the following is true:

- J=l(i.e., vertices represent the same lost pocket from two receivers i and k);
- J € H_k and l € H_i (i.e.,the requested packet of each vertex is in the has set of the receiver that induced the other vertex).

After the construction of the graph, clique partitioning is performed on it. For each clique, a coded packet XORing all the packets are generated and transmitted. Since clique partitioning of a graph is equivalent to the coloring of its complementary graph, the minimum achievable number of re-transmission (T_q) using this technique is equal to

$$T_q = X(G^c)$$

Where $X(G^c)$ is the chromatic number of graph G^c (V,E^c).
After the initial transmission phase, the access node construct a graph G as described, finds a maximal clique in it, and broadcasts an XOR of all the packets represented in its vertices.

Each receiver sends a NAK packet to the access node if it lost this retransmission packet. If there is any loss of packets it will result NAK packets are used by the access node to update the feedback table, which is then used to construct a new graph, and the above mentioned process is re-executed. This process continues until each receiver correctly receives its requested packets.

B. FNCR Scheme

The FNCR scheme has been proposed for packet retransmission to improve wireless multimedia broadcast.

In general, the FNCR scheme combines all the MBS frame packets in each retransmission using linear network coding. Coding coefficients can be either deterministic or selected from a large field such that a large number of coded packets are guaranteed to be linearly independent almost surely. There transmission procedure continues until all receivers get enough packets to decode all packets of the MBS frame.

One drawback of the FNCR scheme for wireless multicast is that it necessitates the delivery of all packets of the MBS frame to all receivers regardless of their needs.

Assuming lossless retransmissions, the number of retransmission packets needed receiver R_i to correctly decoded all the packets is equal to the cardinality of is complimentary set C_i. Consequently, the number of lossless retransmission (T_f) is equal to

$$T_f = \max |C_i| \text{ where } i \epsilon R$$

C. ADAPTIVE NETWORK CODED RETRANSMISSION (ANCR) SCHEME

In this section, the aim to design an efficient and adaptive scheme that can adaptively select the network coded retransmission scheme for each wireless MBS frame. This scheme is refered as the ANCR scheme. The ANCR scheme should select the retransmission scheme that is expected to achieve the smaller number of retransmissions according to the system, demand, and feedback parameters. For the broadcast case, it has been proven that the FNCR scheme is optimal.

Therefore, it will be focusing on the multicast case. For each MBS frame, the ANCR scheme selects one of the two schemes by comparing metrics representing the number of retransmissions for each of them. The scheme having the lower metric is selected to be executed for this MBS frame.

PERFORMANCE ANALYSIS

To estimate their performance it uses two methods.

Method 1: Estimate their performance through their number of lossless retransmission

$$T_q = X(G^c) \text{ and } T_f = \max |C_i| \text{ where } i \epsilon R$$

Since finding the chromatic number of a graph is NP-hard , is need to find an approximation for $X(G^c)$.

METHOD 2: Then extend the lossless ONCR graph representation by including loss pattern information in it, thus generating a lossy ONCR graph model negotiation of G^c. then can estimate the chromatic number of this new graph and compare it to a lossy approximation of the FNCR scheme performance.In both the method the chromatic number of a graph is estimated.

To compare between two methods and decide the complexity level needed to obtain an efficient algorithm, the two approaches are applied to derive π in G^c, then proposing a design for negotiation of G^c and compute its π accordingly.

Approach 1:

In this approach, both the vertices identities and the content of the Has,Complementary and Wants sets of all receivers. This approach consider only the graph vertex set size (V), the system parameters (M,N) and the cardinalities of the different sets that can be extracted from the feedback table. In this approach, the number of retransmissions mostly depends on the cardinalities of the sets and not their contents.

Approach 2:

In this approach, it ignores these cardinalities in addition to the vertices identities. Consequently, this approach considers only the vertex set size V, the system parameters (M,N) and the parameters of the packet request- loss random process. But these approaches are not suite for practices.

Approach 3:

In the previous two approaches, it ignores retransmission packet losses that might occur at different receivers when estimating the ONCR and FNCR performances. In this approach, it aim to consider these loss possibilities in the estimation model to test whether this achieves a better performance than the previous two approaches. Since we don't know the loss realization that will occur during the retransmission phase at the selection time, we will assume that an average number of loss events will occur at each of the receivers.

Then comparing the performance of our three proposed ANCR approaches to both ONCR and FNCR schemes for different number of receivers and demand ratios. As a comparison reference, it defines the optimal selection scheme (denoted as OPT in the figure) as the one that always employs the network coded retransmission scheme that achieves the smaller number of retransmissions.

Performance Result:

For Approaches 1-3, the average and standard deviation of bandwidth efficiency achieved by the ONCR, FNCR, optimal selection, and ANCR schemes against the number of receivers M for N=30 and μ= 0.4. The Metric employed to evaluate different value of this term is the Selection Success Probability. ANCR Scheme succeeds in selecting with lower number of retransmission, divided by the total number of trials.
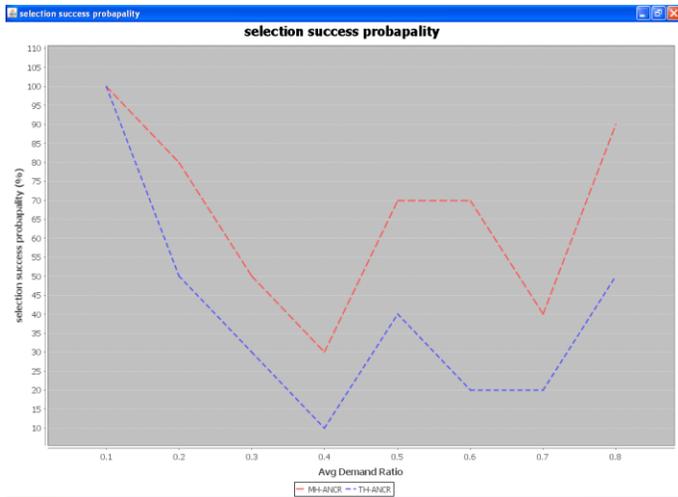
Fig.1. Selection success probablity Multi-hop vs Single hop ANCR
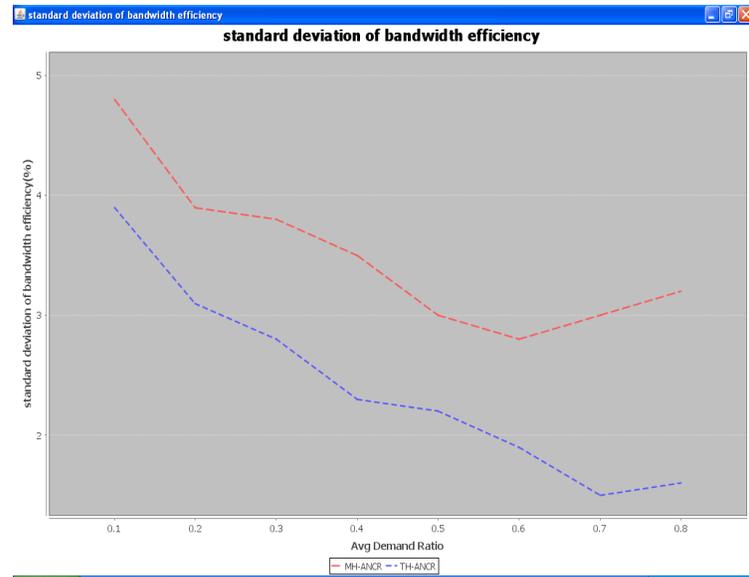
Fig.2. Average bandwidth efficiency Multi-hop vs Single hop ANCR

Fig.3. Standard deviation of bandwidth efficiency Multi-hop vs Single hop ANCR

V. CONCLUSION

In this paper, it has designed an adaptive scheme for packet retransmissions to improve the bandwidth efficiency in wireless MBS using a combination of opportunistic and full network coding. The proposed scheme selects, between these two schemes, the one that is expected to achieve the better bandwidth efficiency performance. To compare between different complexity levels, we presented three selection approaches. In the first two approaches, we derived ONCR performance metrics by modeling its lossless graph representation by a random graph, then using random graph theory. These metrics are then compared to the lossless FNCR performance expression in order to perform scheme selection.

To test the effect of loss patterns on our proposed a third approach in which we first designed a new lossy ONCR graph representation by incorporating an average level of packet losses inside the graph. We then derived a lossy ONCR performance metric that is compared to a lossy approximation of the FNCR performance to perform scheme. For the three considered approaches, simulation results showed that our proposed scheme almost achieves the bandwidth efficiency performance that could be obtained by the optimal selection between the ONCR and FNCR schemes.

REFERENCES

[1] S.Sorour and S.Valaee, "Adaptive network coded retransmission Scheme for wireless multicast," in *Proc. IEEE ISIT*, Jun. 2009, pp.2577–2581.

[2] S. El Rouayheb, M. Chaudhry, and A. Sprintson, "On the minimum Number of transmissions in single-hop wireless coding networks," in *Proc. IEEE ITW*, Sep. 2007, pp. 120–125.

[3] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Wireless broadcasting using network coding," in *Proc. 3rd NetCod*, Jan. 2007, pp. 1–6.

[4] T. Tran, T. Nguyen, and B. Bose, "A joint network-channel coding technique for single-hop wireless networks," in *Proc. 4th NetCod*, Jan.2008, pp. 1–6.

[5] D. Nguyen, T. Tran, T. Nguyen, and B. Bose, "Hybrid ARQ-random network coding for wireless media streaming," in *Proc. 2nd ICCE*, Jun.2008, pp. 115–120.

[6] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[7] B. Bollobas, "The chromatic number of random graphs," *Combinatorica* vol. 8, no. 1, pp. 49–55, Mar. 1988.

[8] M. Chaudhry and A. Sprintson, "Efficient algorithms for index coding," in *Proc. IEEE INFOCOM*, April 2008, pp. 1–4.

[9] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *Proc. ACMSIGCOMM*, 2006, vol. 36, no. 4, pp. 243–254.

[10] J. Sundararajan, D. Shah, and M. Médard, "ARQ for network coding," in *Proc. IEEE ISIT*, Jul. 2008, pp. 1651–1655.

[11] A. Yazdi, S. Sorour, S. Valaee, and R. Kim, "Optimum network coding for delay sensitive applications in wimax unicast," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 2576–2580.

[12] A. Eryilmaz, A. Ozdaglar, M. Médard, and E. Ahmed, "On the delay and throughput gains of coding in unreliable networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5511–5524, Dec. 2008.

[13] D. S. Lun, M. Médard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Phys. Commun.*, vol. 1, no. 1, pp. 3–20, Mar. 2008.

Recognition of Dry and Blurred Fingerprint Using Local Entropy Thresholding Method

T. Ravi[1], B.M.S. Rani[1], Ch. Rajesh Babu[2], Ch. Mounika[2], T. Prashanth[2] K.V.V. Kumar[3], M. Rakesh[3]

[1]Associate professor, Department of ECE, K.L.University, Guntur, A.P, India
[2]B. Tech Student , Department of ECE, K L University, Guntur, A.P, India
[3]M. Tech Student , Department of ECE, K L University, Guntur, A.P, India

ABSTRACT
Dry fingerprint has blurred image, because of overlapping of valleys, singular point to the fingertip and ridges. By using local entropy thresholding method, we extract the fingerprint images. This method is compared to canny filter images, sobel filter images and otsu algorithm images. This technique is used to enhance the contrast between valleys and ridges of the poor, medium and best image features. The experimental results gives the enhancement of the proposed method to get high performance compared with existing techniques. *Keywords - DWT,* Enhancement, Histogram equalization, SVD, Thresholding.

I. INTRODUCTION
wavelet is a wave-like oscillation with an amplitude that starts out at zero, increases, and then decreases back to zero.[1] It can typically be visualized as a "brief oscillation" like one might see recorded by a seismograph or heart monitor. As a mathematical tool, wavelets can be used to extract information from many different kinds of data, including - but certainly not limited to - audio signals and images. Sets of wavelets are generally needed to analyze data fully. A set of "complementary" wavelets will deconstruct data without gaps or overlap so that the deconstruction process is mathematically reversible. Thus, sets of complementary wavelets are useful in wavelet based compression/decompression algorithms where it is desirable to recover the original information with minimal loss. Wavelets are mathematical functions defined over a finite interval and having an average value of zero that transform data into different frequency components, representing each component with a resolution matched to its scale. The basic idea of the wavelet transform is to represent any arbitrary function as a superposition of a set of such wavelets or basis functions. These basis functions or baby wavelets are obtained from a single prototype wavelet called the mother wavelet, by dilations or contractions (scaling) and translations (shifts). They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Many new wavelet applications such as image compression, turbulence, human vision, radar, and earthquake prediction are developed in recent years.

In wavelet transform the basic functions are wavelets. Wavelets tend to be irregular and symmetric. All wavelet functions, $w(2kt - m)$, are derived from a single mother wavelet, $w(t)$.

II. DISCRETE WAVELET TRANSFORM
Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates an awful lot of data. If the scales and positions are chosen based on powers of two, the so-called dyadic scales and positions, then calculating wavelet coefficients are efficient and just as accurate. This is obtained from discrete wavelet transform (DWT).

2.1 2-D WAVELET TRANSFORM HIERARCHY
The 1-D wavelet transform can be extended to a two-dimensional (2-D) wavelet transform using separable wavelet filters. With separable filters the 2-D transform can be computed by applying a 1-D transform to all the rows of the input, and then repeating on all of the columns

LL1	HL1
LH1	HH1

Fig1.1: Subband Labeling Scheme for a one level, 2-D Wavelet Transform

The original image of a one-level (K=1), 2-D wavelet transform, with corresponding notation is shown in Fig. 1. The example is repeated for a three level (K=3) wavelet expansion in Fig. 2. In all of the discussion *K* represents the highest level of the decomposition of the wavelet transform. The 2-D subband decomposition is just an extension of 1-D subband decomposition. The entire process is carried out by executing 1-D subband decomposition twice, first in one

direction (horizontal), then in the orthogonal (vertical) direction. For example, the low-pass subbands (L1) resulting from the horizontal direction is further decomposed in the vertical direction, leading to LL1 and LH1 subbands.

Fig 2.1: Subband labeling Scheme for a Three Level, 2-D Wavelet Transform

Similarly, the high pass subband (Hi) is further decomposed into HLi and HHi. After one level of transform, the image can be further decomposed by applying the 2-D subband decomposition to the existing LL1 subband. This iterative process results in multiple "transform levels". To obtain a two-dimensional wavelet transform, the one-dimensional transform is applied first along the rows and then along the columns to produce four sub bands: low-resolution, horizontal, vertical, and diagonal. (The vertical sub band is created by applying a horizontal high-pass, which yields vertical edges.) At each level, the wavelet transform can be reapplied to the low-resolution sub band to further decorrelate the image.

2.2 HAAR TRANSFORM
In mathematics, the Haar wavelet is a certain sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. Haar used these functions to give an example of a countable orthonormal system for the space of square-integrable functions on the real line. The study of wavelets, and even the term "wavelet", did not come until much later. As a special case of the Daubechies wavelet, it is also known asD2. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of an orthonormal function basis. The Haar sequence is now recognised as the first known wavelet basis and extensively used as a teaching example in the theory of wavelets.

The Haar wavelet's mother wavelet function
    φ(t) can be described as

$$\boldsymbol{\varphi(t)} = \boldsymbol{-1} \quad \frac{1}{2} \le \boldsymbol{t} < 1,$$
$$= 1 \quad 0 \le \boldsymbol{t} < \frac{1}{2},$$
$$= 0 \quad \textbf{otherwise}$$

2.2.1 Propertiesi.
Any continuous real function can be approximated by linear combinations $\phi(t)$, $\phi(2t)$…….$\phi(2^k t)$,…. and their shifted functions. This extends to those function spaces where any function therein can be approximated by continuous functions.

ii. Any continuous real function can be approximated by linear combinations of the constant function, $\varphi(t), \varphi(2t) … … . \varphi(2^k t)$…… and their shifted functions.

III. SVD
In linear algebra, the singular value decomposition (SVD) is a factorization of a real or complex matrix, with many useful applications in signal processing and statistics. Formally, the singular value decomposition of an mxn real or complex matrix M is a factorization of the form
M=UΣV*
where U is an mxm real or complex unitary matrix, Σ is an mxn rectangular diagonal matrix with nonnegative real numbers on the diagonal, and V* (the conjugate transpose of V) is an nxn real or complex unitary matrix[2]. The diagonal entries $\Sigma_{i,i}$ of Σ are known as the singular values of M. The m columns of U and the n columns of V are called the left singular vectors and right singular vectors of M, respectively.

The singular value decomposition and the eigen decomposition are closely related. Namely:

The left singular vectors of M are eigenvectors of MM[*]. The right singular vectors of M are eigenvectors of M[*] M. The non-zero singular values of M (found on the diagonal entries of Σ) are the square roots of the non-zero eigen values of M[*] M or MM[*].

Applications which employ the SVD include computing the pseudo inverse, least squares fitting of data, matrix approximation, and determining the rank, range and null space of a matrix.

IV. FUZZY MEASURES
The discovery of useful information is the essence of any data mining process. Decisions are not usually taken based on complete real world data, but most of the times they deal with uncertainty or lack of information[3]. Therefore the real world reasoning is almost always approximate. However it is not only necessary to learn new information in any data mining process, but it is also important to

understand why and how the information is discovered. Most data mining commercial products are black boxes that do not explain the reasons and methods that have been used to get new information.

However the 'why and how' the information is obtained can be as important as the information on its own. When approximate reasoning is done, measures on fuzzy sets and fuzzy relations can be proposed to provide a lot of information that helps to understand the conclusions of fuzzy inference processes.

Those measures can even help to make decisions that allow to use the most proper methods, logics, operators for connectives and implications, in every approximate reasoning environment. The latest concepts of measures in approximate reasoning is discussed and a few measures on fuzzy sets and fuzzy relations are proposed to be used to understand why the reasoning is working and to make decisions about labels, connectives or implications, and so a few useful measures can help
to have the best performance in approximate reasoning and decision making processes.

Before some measures on fuzzy sets and fuzzy relations are proposed, this chapter collects all the latest new concepts and definitions on measures, and shows a few graphics that make a clear picture on how those measures can be classified. Some important measures on fuzzy sets are the entropy measures and specificity measures. The entropy measures give a degree of fuzziness of a fuzzy set, which can be computed by the premises or outputs of an inference to know an amount of uncertainty crispness in the process. Specificity measures of fuzzy sets give a degree of the utility of information contained in a fuzzy set.

Other important measures can be computed on fuzzy relations. For example, some methods to measure a degree of generalisation of the MODUSPONENS property in fuzzy inference processes are proposed.

4.1 Concept of fuzzy measures
The concept of measure is one of the most important concepts in mathematics, as well as the concept of integral respect to a given measure. The
classical measures are supposed to hold the additive property. Additivity can be very effective and convenient in some applications, but can also be somewhat inadequate in many reasoning environments of the real world as in approximate reasoning, fuzzy logic, artificial intelligence, game theory, decision making, psychology, economy, data mining, etc., that require the definition of non additive measures and a large amount of open problems.

For example, the efficiency of a set of workers is being measured, the efficiency of the same people doing teamwork is not the addition of the efficiency of each individual working on their own.

The concept of fuzzy measure does not require additivity, but it requires monotonicity related to the inclusion of sets. The concept of fuzzy measure

can also be generalized by new concepts of measure that pretend to measure a characteristic not really related with the inclusion of sets. However those new measures can show that "x has a higher degree of a particular quality than y" when x and y are ordered by a preorder (not necessarily the
set inclusion preorder).

The term fuzzy integral uses the concept of fuzzy measure. There are some important fuzzy integrals, as Choquet integral in 1974, which does not require
an additive measure (as Lebesgue integral does).

4.2 Fuzzy set theory
Fuzzy set theory defines set membership as a possibility distribution. The general rule for this can expressed as:

$f : [0,1]^n \rightarrow [0,1]$

where n some number of possibilities. This basically states that we can take n possible events and us f to generate as single possible outcome. This extends set membership since we could have varying definitions of, say, hot curries. One person might declare that only curries of Vindaloo strength or above are hot whilst another might say madras and above are hot[4]. We could allow for these variations definition by allowing both possibilities in fuzzy definitions. Once set membership has been redefined we can develop new logics based on combining of sets etc. and reason effectively.

The membership degree can be expressed by a mathematical function that assigns, to each element in the set, a membership degree between 0 and 1.

The -function is used for modeling the membership degrees. This type of function is suitable to represent the set of bright pixels and is defined as

$$\mu_{AS}(x) = S(x;a,b,c)$$

$$= \begin{cases} 0, & x \leq a \\ 2\left\{\dfrac{(x-a)}{(c-a)}\right\}^2, & a \leq x \leq b \\ 1 - 2\left\{\dfrac{(x-a)}{(c-a)}\right\}^2, & b \leq x \leq c \\ 1, & x \geq c \end{cases}$$

where b=(1/2)(a+c) The -function can be controlled through parameters a and c. Parameter is called the crossover point where $\mu_{AS}(b)=0.5$. The higher the gray level of a pixel (closer to white), the higher membership value and vice versa. A typical shape of the Z-function is presented in Fig. 1.

The -function is used to represent the dark pixels and is defined by an expression obtained from -function as follows:

$$\mu_{AZ}(x) = Z(x;a,b,c) = 1 - S(x;a,b,c)$$

Both membership functions could be seen, simultaneously, in Fig. 2. The -function in the right side of the histogram and the Z-function in the left.
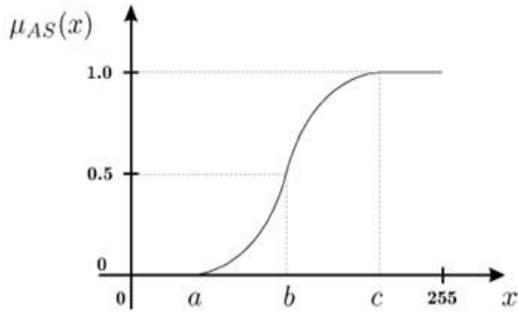

Fig3.1: Typical shape of the S-function
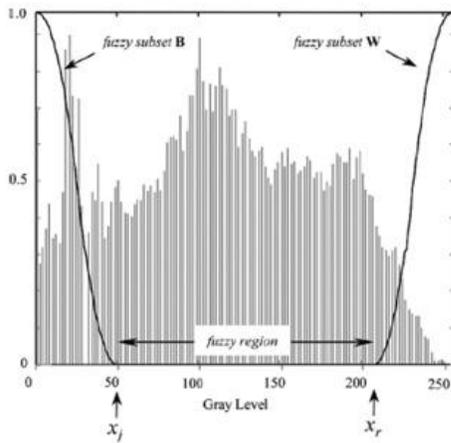
Fig3.2: Histogram and functions for the seed subsets

V. MATLAB RESULTS AND GRAPHS

In this section, simulation results for different images (64x64) are shown. Their histogram measure graphs are also included. Considered the images fingerprint.jpeg form the MATLAB library

Fig 4.1:original Image

Fig 4.2 :Red Image
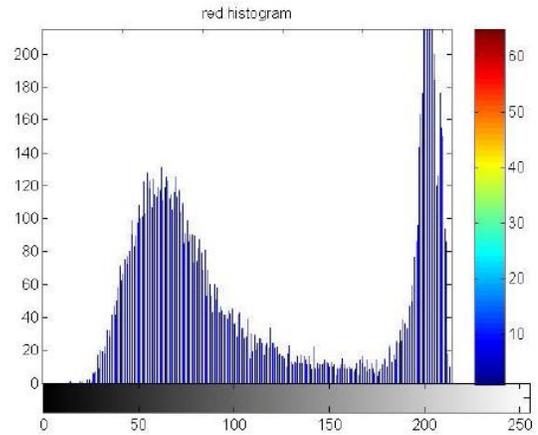
Fig 4.3: Red histogram
Histogram measuring pixel values
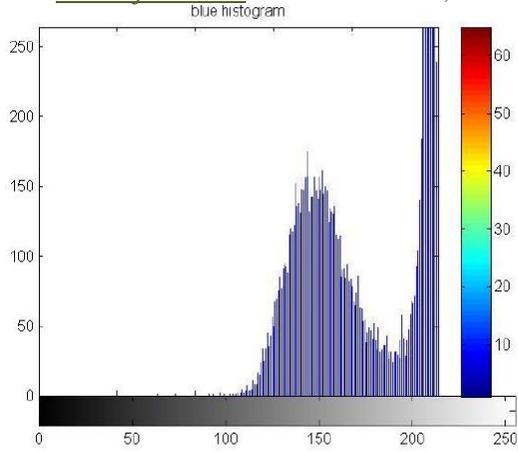
Fig 4.4: Blue Image

Fig 4.5:Blue histogram

Fig4.6 Green Image
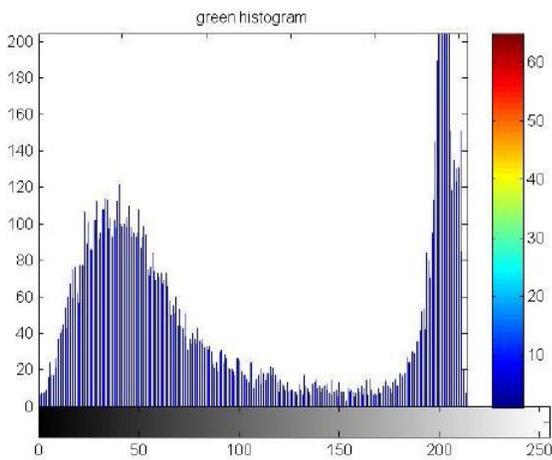
Fig 4.7:Green histogram

Fig 4.8: Histogram of green and red
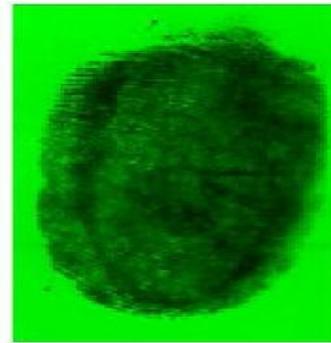
Fig 4.9:Histogram of green and blue
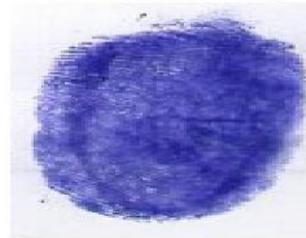
Fig 4.10:Histogram of red and blue

Fig 4.11: Canny Image

Fig 4.12: Sobel Image

Fig 4.12 Prewitt Image

5. CONCLUSION

Our method gives the better performance, compare with existing algorithms. it is applicable for Dry finger images and blurred images.

REFERENCES

[1]. The Discrete Wavelet Transform: Algorithms And Application Mark By D. Konezny

[2]. Projection Matrices, Generalized Inverse Matrices, And Singular Value Decomposition by Haruo Yanai, Kei Takeuchi, Yoshio Takane

[3]. Introduction To Neural Networks Using MATLAB 6.0 (Computer Science Series) by S Sivanandam, S Sumathi.

[4]. FUZZY SET THEORY: BASIC CONCEPTS, TECHNIQUES AND BIBLIOGRAPHY R LOWEN.

[5]. Face Recognition Techniques and Analysis: Classification in Principal and Histogram Spaces by Kyperountas (Jul 3, 2009).

[6]. An Introduction to Wavelet Analysis by David F. Walnut (Sep 27, 2001).

[7]. Fuzzy Min-Max Neural Networks- Part1 classification by Patrick k Simson (September 1992).

[8]. Neural Networks Based Fuzzy Logic Control and Decision System by C. S. George Lee (December 1991).

[9]. Illuminant and Device Invariant Color Using Histogram Equalisation by Elsevier Computer Science (December 2002).

[10]. Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization By Yeong-Taeg Kim

Author's Bibliography

K. V. V. Kumar was born in 1987 at Guntur district of Andhra Pradesh state, india. He Graduated in Electronics and Communication Engineering from VYCET, JNTU, Kakinada. Presently he is pursuing his M.Tech-Communications and Radar Systems in K. L. University. His interested areas are Image Processing, Antennas and Communications.

M. Rakesh was born in 1987 at Guntur district of Andhra Pradesh state, India. He Graduated in Electronics and Communication Engineering from KMCET, JNTU, Hyderabad. Presently he is pursuing his M. Tech-Communications and Radar Systems in K. L. University. His interested areas are Image Processing and Communications.

Securing Layer-3 Wormhole Attacks in Ad-Hoc Networks

T. Krishna Rao[1]
[1]PG Scholar,
Department of IT,
Aurora Engineering College,
Bhongir, Nalgonda Dist,
Andhra Pradesh, India

Mayank Sharma[2]
[2]Associate Professor
Department of IT
Aurora Engineering College
Bhongir ,Nalgonda Dist,
Andhra Pradesh , India

Dr. M. V. Vijaya Saradhi[3]
[3]Professor, Head of the Department
Department of IT
Aurora Engineering College
Bhongir, Nalgonda Dist,
Andhra Pradesh,India

Abstract - In ad hoc networks, malicious nodes can carry wormhole attacks to fabricate a false scenario on neighbour relations among mobile nodes. The attacks threaten the safety of ad hoc routing protocols and some security enhancements. In the wormhole attack, an attacker records packets (or bits) at one location in the network, tunnels them (possibly selectively) to another location, and retransmits them into the network. The wormhole attack can form a serious threat in wireless networks, especially against many ad hoc network routing protocols and location-based wireless security systems. In this paper, we present a new approach for detecting wormhole attacks. The Witness Integration Multipath protocol is based on the Multipath DSR routing protocol and finds suspicious behavior related to wormhole attacks.

Keywords: - MANET, source routing, multipath, wormhole attacks

I. INTRODUCTION

Ad-hoc networks must deal with threats from external agents and compromised internal nodes. The lack of a central control and the fact that each node must forward packets of other nodes represent major security challenges. In such environments, it is difficult to assure the confidentiality and the integrity of the communications as well as the availability of the services. Mobile ad-hoc networks [1]have been an attractive field of research for many years now. Due to their characteristics, these networks are an excellent choice for emergency operations, vehicular communication and short-live networks.

In the Wormhole attack, an attacker records a packet or individual bits from a packet, at one location in the network, tunnels the packet (possibly selectively) to another location, and replays it there. It is simple for the attacker to make the tunnelled packet arrive with better metric than a normal multihop route. This can be done for tunnelled distances longer than the normal wireless transmission range of a single hop. It is also possible for the attacker to forward each bit over the wormhole directly, without waiting for an entire packet to be received before beginning to tunnel the bits of the packet, in order to minimize delay introduced by the wormhole. Due to the nature of wireless transmission, the attacker can create a wormhole even for packets not addressed to it, since it can overhear them in wireless transmission and tunnel them to the colluding attacker at the opposite end of the wormhole.

The paper is organized as follows. Section 2 presents an Overview of Wormhole Attacks. Section 3 discusses the related works. Section 4 presents the approach. Section 5 gives an overview of implementation and results and Section 6 presents the conclusions.

II. OVERVIEW OF WORMHOLE ATTACKS

In a *wormhole attack*, an attacker receives packets at one point in the network, "tunnels" them to another point in the network, and then replays them into the network from that point. For tunnelled distances longer than the normal wireless transmission range of a single hop, it is simple for the attacker to make the tunnelled packet arrive with better metric than a normal multihop route, for example through use of a single long-range directional wireless link or through a direct wired link to a colluding attacker. It is also possible for the attacker to forward each bit over the wormhole directly, without waiting for an entire packet to be received before beginning to tunnel the bits of the packet, in order to minimize delay introduced by the wormhole. Due to the nature of wireless transmission, the attacker can create a wormhole even for packets not addressed to it, since it can overhear them in wireless transmission and tunnel them to the colluding attacker at the opposite end of the wormhole[2].

If the attacker performs this tunnelling honestly and reliably, no harm is done; the attacker actually provides a useful service in connecting the network more efficiently. However, the wormhole puts the attacker in a very powerful position relative to other nodes in the network, and the attacker could exploit this position in a variety of ways. The attack can also still be performed even if the network communication provides confidentiality and authenticity, and even if the attacker has no cryptographic keys. Furthermore, the attacker is invisible at higher layers; unlike a malicious node in a routing protocol, which can often easily be named, the presence of the wormhole and the two colluding attackers at either endpoint of the wormhole are not visible in the route.
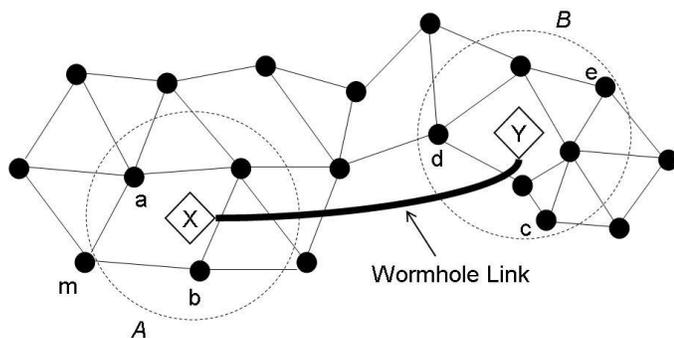
Figure 1: Wormhole Attack

An example of Wormhole attack is shown in the figure 1. Here X and Y are the two end-points of the wormhole link (called as wormholes)[4]. X replays in its neighborhood (in area A) everything that Y hears in its own neighborhood (area B) and vice versa. The net effect of such an attack is that all the nodes in area A assume that nodes in area B are their neighbors and vice versa. This, as a result, affects routing and other connectivity based protocols in the network. Once the new routes are established and the traffic in the network starts using the X-Y shortcut, the wormhole nodes can start dropping packets and cause network disruption. They can also spy on the packets going through and use the large amount of collected information to break any network security. The wormhole attack will also affect connectivity-based localization algorithms and protocols based on localization, like geographic routing, will find many inconsistencies resulting in further network disruption.

The wormhole attack is particularly dangerous against many ad hoc network routing protocols in which the nodes that hear a packet transmission directly from some node consider themselves to be in range of (and thus a neighbour of) that node. For example, when used against an on-demand routing protocol such as DSR [16], [17] or AODV [27], a powerful application of the wormhole attack can be mounted by tunnelling each ROUTE REQUEST packet directly to the destination target node of the REQUEST. When the destination node's neighbours hear this REQUEST packet, they will follow normal routing protocol

processing to rebroadcast that copy of the REQUEST and then discard without processing all other received ROUTE REQUEST packets originating from this same Route Discovery. This attack thus prevents any routes other than through the wormhole from being discovered, and if the attacker is near the initiator of the Route Discovery, this attack can even prevent routes more than two hops long from being discovered. Possible ways for the attacker to then exploit the wormhole include discarding rather than forwarding all data packets, thereby creating a permanent Denial-of-Service attack (no other route to the destination can be discovered as long as the attacker maintains the wormhole for ROUTE REQUEST packets), or selectively discarding or modifying certain data packets.

The neighbour discovery mechanisms of periodic (proactive) routing protocols such as DSDV [26], OLSR [33], and TBRPF [5] rely heavily on the reception of broadcast packets as a means for neighbour detection, and are also extremely vulnerable to this attack. For example, OLSR and TBRPF use HELLO packets for neighbour detection, so if an attacker tunnels through a wormhole to a colluding attacker near node B all HELLO packets transmitted by node A, and likewise tunnels back to the first attacker all HELLO packets transmitted by B, then A and B will believe that they are neighbours, which would cause the routing protocol to fail to find routes when they are not actually neighbours. For DSDV[3], if each routing advertisement sent by node A or node B were tunnelled through a wormhole between colluding attackers near these nodes, as described above, then A and B would believe that they were neighbours. If A and B, however, were not within wireless transmission range of each other, they would be unable to communicate. Furthermore, if the best existing route from A to B were at least $2n + 2$ hops long, then any node within n hops of A would be unable to communicate with B, and any node within n hops of B would be unable to communicate with A. Otherwise, suppose C were within n hops of A, but had a valid route to B. Since A advertises a metric of 1 route to B, C would hear a metric n+1 route to B. C will use that route if it is not within n+1 hops of B, in which case there would be an n-hop route from A to C, and a route of length n+1 from C to B, contradicting the premise that the best real route from A to B is at least $2n + 2$ hops long.

In each of these protocols, the wormhole can be used to attract ad hoc network traffic, and can use this position to eavesdrop on traffic, maliciously drop packets, or to perform man-in-the-middle attacks against protocols used in the network. The wormhole attack is also dangerous in other types of wireless networks and applications. One example is any wireless access control system that is based on physical proximity, such as wireless car keys, or proximity and token based access control systems for PCs [8], [20]. In such systems, an attacker could relay the authentication exchanges to gain unauthorized access.

III. RELATED WORKS

Due to the characteristics of the wormhole attacks, cryptographic solutions are not sufficient. Numerous physical approaches have been proposed to secure the neighbor discovery process. Most of the solutions presented so far require that the nodes handle information about self-location, perform clocks synchronization or rely on specialized antennas or on information such as trust relationship. Only few solutions have been proposed to secure the overall end-to-end route discovery process. Other approach contains timing and/or position information to packets. This restricts the maximum transmission distance permitted to a packet. They propose two kinds: geographical and temporal. To use geographical approach, each node must know its own location and all nodes must have loosely synchronized clocks. To use temporal approach, all nodes must have tightly synchronized clocks. Thus, if a receiving node determines that the neighbor discovery signal of a given node has traveled too far, the node should discard it. Another approach is to estimate the distance separating two nodes from the round-trip travel time taken by a message and its acknowledgement. This mechanism relies on a specialized hardware allowing the destination to send a response to a one bit challenge message as fast as possible.

Several approaches have been developed to prevent or to detect wormhole attacks. The first three solutions address mainly the closed wormhole attacks. They present how to protect the neighbour discovery process. Hu *et al.* [15] propose the addition of *leashes* containing timing and/or position information to packets. A leash restricts the maximum transmission distance permitted to a packet. They propose two kinds of leashes: geographical and temporal. To use geographical leashes, each node must know its own location (e.g. GPS) and all nodes must have loosely synchronized clocks. To use temporal leashes, all nodes must have tightly synchronized clocks. Thus, if a receiving node determines that the neighbour discovery beacon of a given node has travelled too far, the node should discard it.

Cˇ apkun *et al.* [16] estimates the distance separating two nodes from the round-trip travel time taken by a message and its acknowledgement. This mechanism relies on a specialized hardware allowing the destination to send a response to a one bit challenge message as fast as possible. Hu and Evans [17] use directional antennas to detect wormhole attacks. If a node uses a specific sector to communicate with a neighbour, this neighbour should use its opposite sector. The existence of a wormhole would introduce inconsistencies in the network that could be detected by the other nodes simply by adding some sector information to the packets. The next solutions address the open wormhole attacks. They present how to prevent or detect malicious actions from compromised internal agents. Pirzada and McDonald [18] derive a trust relationship for neighbour nodes based upon their compliance to a routing protocol (DSR). The nodes' trust levels are then used to avoid communication through potential wormholes.

Khalil *et al.* [18] propose that the nodes in a static network obtain in a secure way the one-hop and two-hop topological information from their neighbours. Then, each node observes the behaviour of their neighbours searching for typical patterns related to wormhole attacks. The same authors also propose to support nodes mobility by adding a trusted central authority in charge of authorizing nodes to move and to create new neighbour associations [20]. Wang *et al.* [14] extend the geographical leashes and use them in an end-to-end verification process. This process determines whether all the supposedly neighbour pairs of a path are not too far apart. Finally, Qian *et al.* [21] present a different approach to detect wormhole attacks. The solution is based on statistical analysis of the information gathered during the multipath routing process (SMR). A link generating a wormhole attack should be used by the routing protocol with an unusually high frequency. Unfortunately, only uniform grid networks have been considered.

IV. APPROACH

The Dynamic Source Routing (DSR) protocol [22] is an on demand source routing protocol for mobile ad-hoc networks. When a source needs a path towards a destination, it broadcasts *Route Request* (RREQ) messages. As these messages are forwarded, they gather the intermediate nodes they go through. Then, the destination replies with unicast *Route Reply* (RREP) messages to the source. The source chooses its path based on the received RREP messages. To avoid too many RREQ packets in the network, the protocol uses two mechanisms: local cache and selective broadcasting for intermediate nodes. An intermediate node can respond if it has a valid path in its cache. Otherwise, it forwards the request message if it is a new one.

The DSR protocol has been adapted to discover disjoint multipath between a source and a destination. Using multiple paths can improve the quality of service as well as the fault resilience of a network. The routing protocol used in this paper is based on a modification of the Split Multipath Routing (SMR) protocol [23] proposed by Quian *et al.* [21]. The modified protocol allows intermediate nodes to forward repeated copies of a RREQ message, as long as their hop counts are not larger than the hop counts of already received copies. The destination should receive numerous copies of the RREQ message. Thus, the destination should be able to build a list of available paths from the source; this information gives a partial view of the network that would be used by the WIM-DSR protocol in the discovery of possible wormhole attacks.

The WIM-DSR final step is slightly different from the previous protocols. The main objective of WIM-DSR is to gather information during the route discovery phase and to find possible anomalies due to open wormhole attacks The

destination chooses a path and broadcasts it towards the source. The intermediate nodes should rebroadcast only one copy of a given RREQ message. This step should allow intermediate nodes to validate the information. WIM-DSR determines if the information gathered by the modified routing protocol during the route discovery shows the typical behaviour of wormhole attacks.

The aim of WIM-DSR is to find *fully witnessed* paths, i.e. paths with only witnessed edges between the source and the destination. Fully witnessed path should not contain any open wormhole. Strongly witnessed paths should be preferred. However, weakly witnessed paths should also be considered since the strongly witnessed condition is very restrictive and can generate numerous false positive alarms.

Once a fully witnessed path is found, the destination signs its RREP message and broadcasts it towards the source. For a strong witnessed path, the destination broadcasts a unique signed RREP message which is rebroadcast by all the nodes of the path. The other nodes simply overhear it. This allows each witness to receive the message from at least two nodes. For a path of length l, only $l - 1$ RREP messages are sent overall. For a weak witnessed path, the destination unicasts a signed RREP message along the path. Moreover, for each witness, the destination also unicasts a signed confirmation RREP message along a path going through that witness.

The real gain for the malicious nodes is the strong open wormhole attack. In such a case, they would be selected by any protocol selecting the shortest paths. Such a wormhole represents a shortcut in the network. The effectiveness of WIM-DSR to detect open wormhole attacks is proven in the following lemmas. They show that the path selection algorithms cannot find *false* witnesses for *strong open wormholes*.

V. EVALUATION

The objective is to determine how many pairs of source and destination nodes do not have fully witnessed paths in a given set of points. These pairs would represent the false positive alarms for the protocol. The network density is important for ad-hoc networks. For a given region, there are two ways to increase the density: (1) increase the number of nodes or (2) increase the transmission range of the nodes. Since the complexity of the simulation program depends on the number of nodes, the number of nodes is fixed and different range values are used. The concept of this paper is implemented and different results are shown below.

The proposed approach is implemented in Java and J2ME technology on a Pentium-III PC with 20 GB hard-disk and 256 MB RAM. The propose approach's concepts show efficient results of retrieving data from mobile nodes and has been efficiently tested on different systems.

Figure 2: Sending message
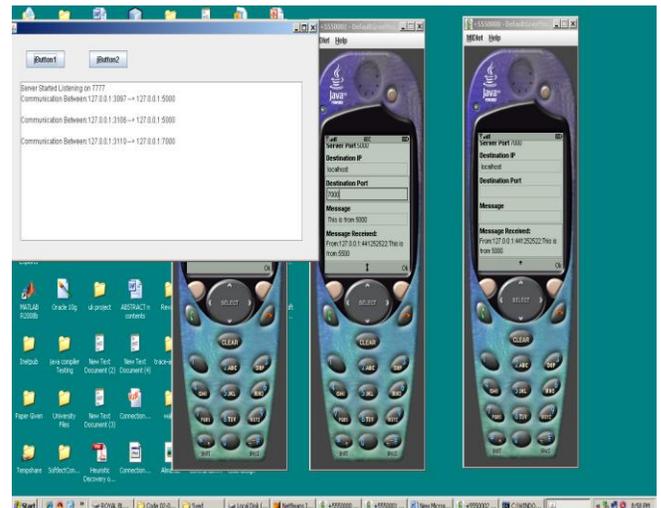
Figure 3: Edge Witnessing

Figure 4: Finding Witnessed paths and detecting attacks

VI. CONCLUSTIONS AND FUTURE ENHANCEMENTS

We introduce a new approach preventing strong open wormhole attacks. The WIM-DSR protocol uses the information collected by the destination node during the route discovery process of a multipath routing protocol to detect suspicious behaviour. The results obtained in this project show that the WIM-DSR protocol is able to detect all strong open wormhole attacks with a very low rate of false positive alarms. This solution does not require any cryptographic processing by the intermediate nodes, if no attack takes place.

Our implementation has shown that the prevention can be considered as reliable. The routing metric packet delivery is high mobility. Characteristics and results for this system were achieved after an extensive design part in our implementation. Design has been a key part to get reliable results. This study could be continued by, for instance, developing the multipath aspect of our protocol. It could be achieved by splitting data packets from the source to the destination; the whole message would not be transmitted by the same path or the same nodes all the time. Another solution could be to enforce reliability adding some redundancy code; in that case, it would allow not sending again packets in case one link breaks.

REFERENCES

[1] Y.-C. Hu, A. Perrig, and D. Johnson, "Wormhole attacks in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 370 – 380, 2006.

[2] "Mobiworp: Mitigation of the wormhole attack in mobile multihop wireless networks," *Ad Hoc Networks*, vol. 6, pp. 344–362, 2008.

[3] D. Johnson and D. Maltz, *Dynamic Source Routing in Ad Hoc Wireless Networks*. Kluwer Academic Publishers, ch. 5, pp. 153 – 181.

[4] S. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," in *Proc. of the IEEE International Conference on Communications*, 2001, pp. 3201 – 3205.

[5] Y.-C. Hu, A. Perrig, and D. Johnson, "Ariadne: A secure on-demand routing protocol for ad hoc networks," *Wireless Networks*, vol. 11, pp. 21 – 38, 2005.

[6] David B. Johnson and David A. Maltz. Dynamic Source Routing in Ad Hoc Wireless Networks. In *Mobile Computing*, edited by Tomasz Imielinski and Hank Korth, chapter 5, pages 153–181. Kluwer Academic Publishers, 1996.

[7] David B. Johnson, David A. Maltz, and Josh Broch. The Dynamic Source Routing Protocol for Multihop Wireless Ad Hoc Networks. In *Ad Hoc Networking*, edited by Charles E. Perkins, chapter 5, pages 139– 172. Addison-Wesley, 2001.

[8] Adrian Perrig, Ran Canetti, Doug Tygar, and Dawn Song. Efficient Authentication and Signature of Multicast Streams over Lossy Channels. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, pages 56–73, May 2000.

[9] Charles E. Perkins and Pravin Bhagwat. Highly Dynamic Destination- Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In *Proceedings of the SIGCOMM'94 Conference on Communications Architectures, Protocols and Applications*, pages 234–244, August 1994.

[10] Amir Qayyum, Laurent Viennot, and Anis Laouiti. Multipoint Relaying: An Efficient Technique for flooding in Mobile Wireless Networks. Technical Report Research Report RR-3898, Project HIPERCOM, INRIA, February 2000.

[11] Bhargav Bellur and Richard G. Ogier. A Reliable, Efficient Topology Broadcast Protocol for Dynamic Networks. In *Proceedings of the Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'99)*, pages 178–186, March 1999

[12] Mark Corner and Brian Noble. Zero-Interaction Authentication. In *Proceedings of the Eighth Annual International Conference on Mobile Computing and Networking (MobiCom 2002)*, pages 1–11, September 2002.

[13] Tim Kindberg, Kan Zhang, and Narendar Shankar. Context Authentication Using Constrained Channels. In *Proceedings of the Fourth IEEE Workshop on Mobile Computing Systems and Applications (WMCSA 2002)*, pages 14–21, June 2002.

[14] W. Wang, B. Bhargava, Y. Lu, and X. Wu, "Defending against wormhole attacks in mobile ad hoc networks," *Wireless Communications and Mobile Computing*, vol. 6, pp. 483 – 502, 2006.

[15] L. Butty´an and J.-P. Hubaux, *Security and Cooperation in Wireless Networks*. Cambridge University Press, 2008.

[16] Y.-C. Hu, A. Perrig, and D. Johnson, "Wormhole attacks in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 370 – 380, 2006.

[17] S. Capkun, L. Butty´an, and J.-P. Hubaux, "Sector: secure tracking of node encounters in multi-hop wireless networks," in *Proc. of the 1st ACM workshop on Security of ad hoc and sensor networks (SASN)*, 2003, pp. 21 – 32.

[18] L. Hu and D. Evans, "Using directional antennas to prevent wormhole attacks," in *Proc. of the Network and Distributed System Security Symposium*, 2004.

[19] A. A. Pirzada and C. McDonald, "Detecting and evading wormholes in mobile ad-hoc wireless networks," *Int. Journal of Network Security*, vol. 3, pp. 191 – 202, 2006.

[20] I. Khalil, S. Bagchi, and N. B. Shroff, "Liteworp: Detection and isolation of the wormhole attack in static multihop wireless networks," *Computer Networks*, vol. 51, pp. 3750–3772, 2007.

[21] "Mobiworp: Mitigation of the wormhole attack in mobile multihop wireless networks," *Ad Hoc Networks*, vol. 6, pp. 344–362, 2008.

[22] L. Qian, N. Song, and X. Li, "Detection of wormhole attacks in multipath routed wireless ad hoc networks: a statistical analysis approach," *Journal of Network and Computer Applications*, vol. 30, pp. 308 – 330, 2007.

[23] D. Johnson and D. Maltz, *Dynamic Source Routing in Ad Hoc Wireless Networks*. Kluwer Academic Publishers, ch. 5, pp. 153 – 181.

[24] S. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," in *Proc. of the IEEE International Conference on Communications*, 2001, pp. 3201 – 3205.

Study & Simulation of O.F.D.M System

Ashok Kamboj[1] Geeta Kaushik[2]
ECE Deptt.M.M.University Mullana (India)

ABSTRACT
In this paper, we describe the Orthogonal frequency division multiplexing (OFDM) is an established technique for wireless communication applications. Typical constraints faced during OFDM transmission are: a large peak-to-mean envelope power ratio, which can result in significant distortion when transmitted through a nonlinear device, such as a transmitter power amplifier. We study the effects of clipping and filtering on the performance of OFDM, including the power spectral density, BER, through intensive MATLAB simulation. We have indigenously simulated the effect of multipath fading to ensure that all specifications of OFDM transmission are taken care of. To simulate the modulation of the sub-carriers, we have chosen DQPSK. The way OFDM handles ISI has also been encompassed.

General Terms
Orthogonal frequency division multiplexing (OFDM) for wireless communication.

1. INTRODUCTION
It is valuable to discuss the mathematical definition of OFDM. This allows us to see how the signal is generated and how the receiver must operate, and it gives us a tool to understand the effects of imperfections in the transmission channel. OFDM transmits a large number of narrowband carriers, closely spaced in the frequency domain. In order to avoid a large number of modulators and filters at the transmitter and complementary filters and demodulators at the receiver, it is desirable to be able to use modern digital signal processing techniques, such as fast Fourier transform (FFT) [2]. Mathematically, each carrier can be described as a complex wave

$$s_c(t) = A_c(t)e^{j\left[\omega_c t + \phi_c(t)\right]}$$
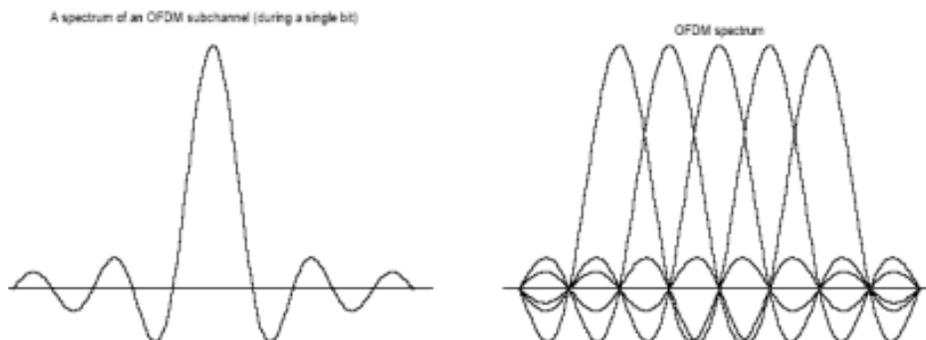
The real signal is the real part of Sc*(t)*. Both *Ac(t)* and *fc(t),* the amplitude and phase of the carrier, can vary on a symbol by symbol basis. The values of the parameters are constant over the symbol duration procedures. The IFFT & FFT operations are naturally performed through easily available commands in MATLABTM

2. OFDM MODEL
All The OFDM system was modeled using MATLABTM to allow various parameters of the system to be varied and tested. The aim of doing the simulations was to measure the performance of OFDM under different channel conditions, and to allow for different OFDM configurations to be tested. Four main criteria were used to assess the performance of the OFDM system, which were its tolerance to multipath delay spread, peak power clipping, channel noise and time synchronization errors.

The OFDM system was modeled using the Communications Toolbox, Signal Processing Toolbox and Simulink of MATLABTM, and is shown in Figure 2.1 A brief description of the model is provided below
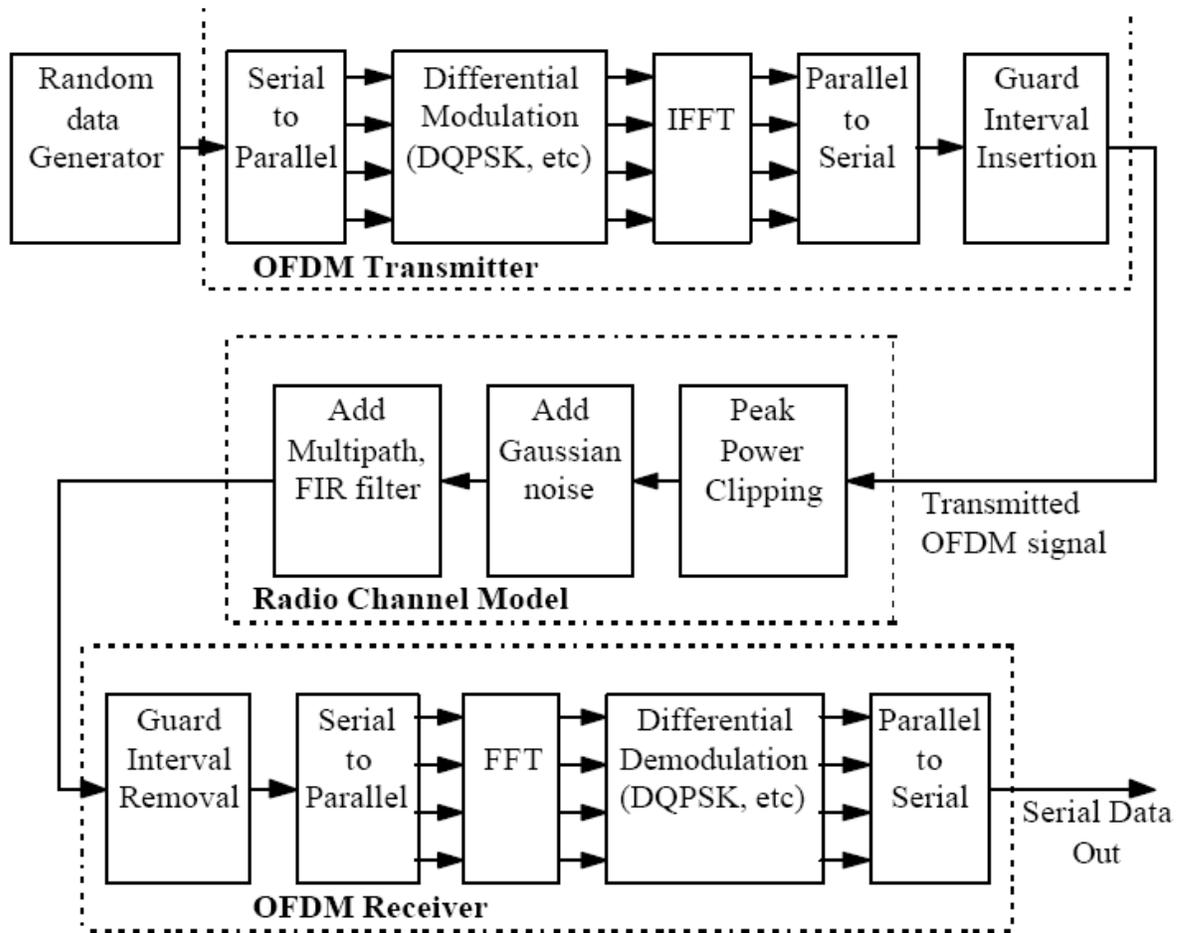
2.1   Fig. OFDM model used for simulation

3. OFDM

3.1 Serial to Parallel Conversion
The input serial data stream is formatted into the word size required for transmission, e.g. 2 bits/word for QPSK, and shifted into a parallel format. The data is then transmitted in parallel by assigning each data word to one carrier in the transmission.

3.2 Modulation of Data
The data to be transmitted on each carrier is then differentially encoded with previous symbols, and then mapped into a PSK format. Since differential encoding requires an initial phase reference an extra symbol is added at the start for this purpose. The data on each symbol is then mapped to a phase angle based on the modulation method. For example, for QPSK the phase angles used are 0, 90, 180, and 270 degrees. The use of phase shift keying produces a constant amplitude signal and was chosen for its simplicity and to reduce problems with amplitude fluctuations due to fading.

3.3 Inverse Fourier Transform
After the required spectrum is worked out, an inverse Fourier transform is used to find the corresponding time waveform. The guard period is then added to the start of each symbol.

3.4 Guard Period
The guard period used was made up of two sections. Half of the guard period time is a zero amplitude transmission. The other half of the guard period is a cyclic extension of the symbol to be transmitted. This was to allow for symbol timing to be easily recovered by envelope detection. However it was found that it was not required in any of the simulations as the timing could be

accurately determined position of the samples. After the guard has been added, the symbols are then converted back to a serial time waveform. This is then the base band signal for the OFDM transmission.

3.5 Channel

A channel model is then applied to the transmitted signal. The model allows for the signal to noise ratio, multipath, and peak power clipping to be controlled. The signal to noise ratio is set by adding a known amount of white noise to the transmitted signal. Multipath delay spread is then added by simulating the delay spread using an FIR filter. The length of the FIR filter represents the maximum delay spread, while the coefficient amplitude represents the reflected signal magnitude.

3.6 Receiver

The receiver basically does the reverse operation to the transmitter. The guard period is removed. The FFT of each symbol is then performed to find the original transmitted spectrum. The phase angle of each transmission carrier is then evaluated and converted back to the data word by demodulating the received phase. The data words are then combined back to the same word size as the original data.

3.7 OFDM generation

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 first choosing the spectrum required, based on the input data, and modulation scheme used. Each carrier to be produced is assigned some data to transmit. The required amplitude and phase of the carrier is then calculated based on DQPSK. The required spectrum is then converted back to its time domain signal using an Inverse Fourier Transform. In most applications, an Inverse Fast Fourier Transform (IFFT) is used. The IFFT performs the transformation very efficiently, and provides a simple way of ensuring the carrier signals produced are orthogonal.

3.8 Adding a Guard Period to OFDM

One of the most important properties of OFDM transmissions is its high level of robustness against multipath delay spread. This is a result of the long symbol period used, which minimizes the inter-symbol interference. The level of multipath robustness can be further increased by the addition of a guard period between transmitted symbols. The guard period allows time for multipath signals from the previous symbol to die away before the information from the current symbol is gathered. The most effective guard period to use is a cyclic extension of the symbol.

If a mirror in time, of the end of the symbol waveform is put at the start of the symbol as the guard period, this effectively extends the length of the symbol, while maintaining the orthogonality of the waveform. Using this cyclic extended symbol the samples required for performing the FFT (to decode the symbol), can be taken anywhere over the length of the symbol. This provides multipath immunity as well as symbol time synchronization tolerance.

4. Parameters of an actual OFDM System

Following are the parameters of Wi-Fi / IEEE 802.11a which is a system based on OFDM:

- Data Rates   :   6 Mbps to 48 Mbps
- Modulation   :   BPSK, QPSK, 16 QAM and 64 QAM
- Coding   :   Convolution concatenated with Reed Solomon
- FFT Size   :   64 with 52 sub-carriers uses, 48 for data and 4 pilots.
- Subcarrier frequency spacing   :   20 MHz divided by 64 carriers or 0.3125 MHz
- FFT Period   :   Also called symbol period, 3.2 µs
- Guard Duration   :   One quarter of symbol time, 0.8 µs
- Symbol Time   :   4 µs

5. IMPLEMENTATION

Implementation refers to the process that bridges the gap between study and simulation. To simulate a system, one needs to obtain some model which can be transferred into the software being used. The model can either be a mathematical model or circuit description or algorithm or anything else which can define the system. The software that we decided to use was Simulink. This software provides a block-diagram approach to simulate systems. Our research/study had also yielded a variety of block diagrams for the OFDM transmitter and receiver. So, the task of implementation became quite easy. The only major task was to search out which blocks were available in the Simulink library that could be used by us. Also, if some blocks were not available, some other blocks with little modifications in their parameters could be used. One may also obtain some blocks by a combination of blocks available in the library. If in case, any of the above mentioned option is not possible, S-functions could be used to model the blocks required.

However, in our implementation process, we did not have to resort the use of S-functions as all the blocks we required were either available directly in the library or were derived from some other blocks available in the library.

Various blocks like QPSK modulators/demodulators, zero pads, AWGN channel, linear binary encoder/decoder, convolution coder/decoder, FFT, IFFT, random binary generator, etc. were available in the Simulink library. For the S/P and P/S blocks we used Buffers. For addition and removal of cyclic prefix, we used Selectors. In making the bock which would add pilots, we used Multiport selector, PN sequence generator and Matrix concatenation blocks. For providing multipath in the channel, we used delay elements and adders.

5. Simulation

Simulation of basic OFDM system was studied. In the study the following features have been taken care of

- Encoding (Linear and Convolution)
- Pilot carriers
- Multipath in the channel

For each simulation, observed the following things:

- BER between the transmitted and received signals
- Constellation diagrams for transmitted and received signals
- Frequency spectrum of the channel

5.1 Simulation Results for OFDM with pilots, multipath and codes (convolutional)
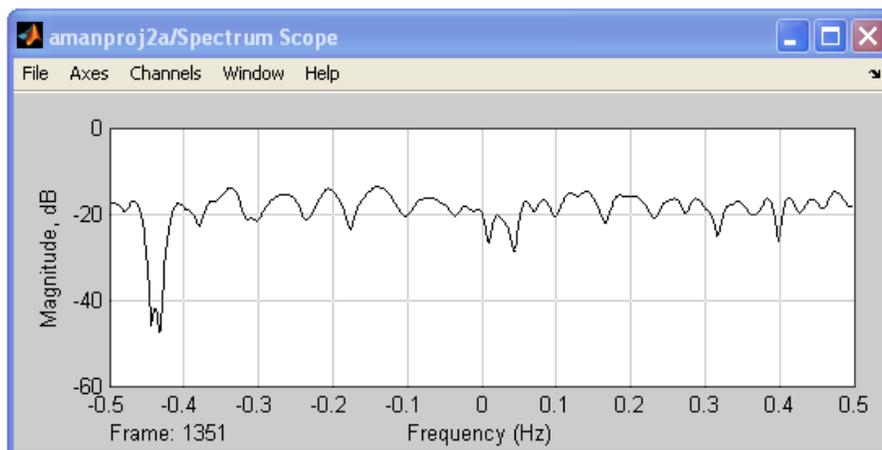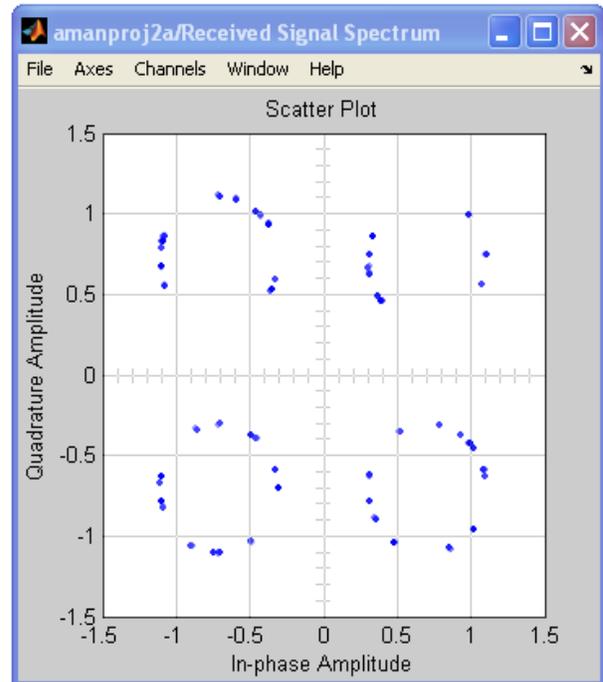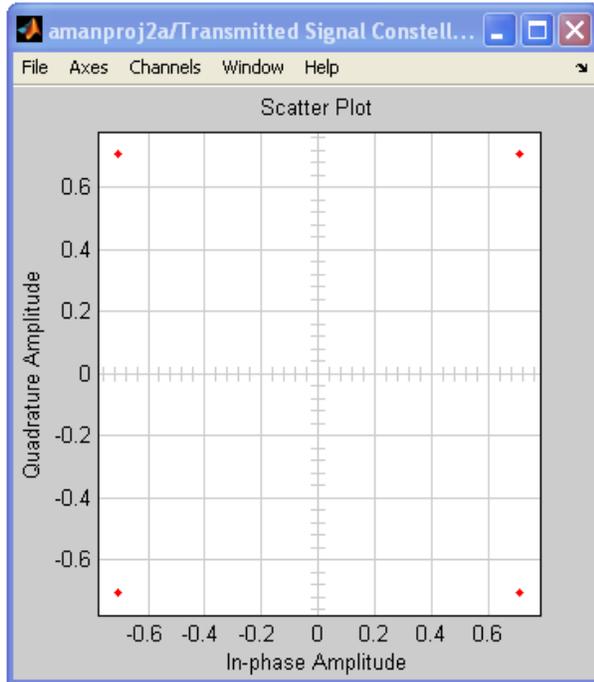
5.2 Simulation Results for OFDM with pilots, multipath and codes (linear)
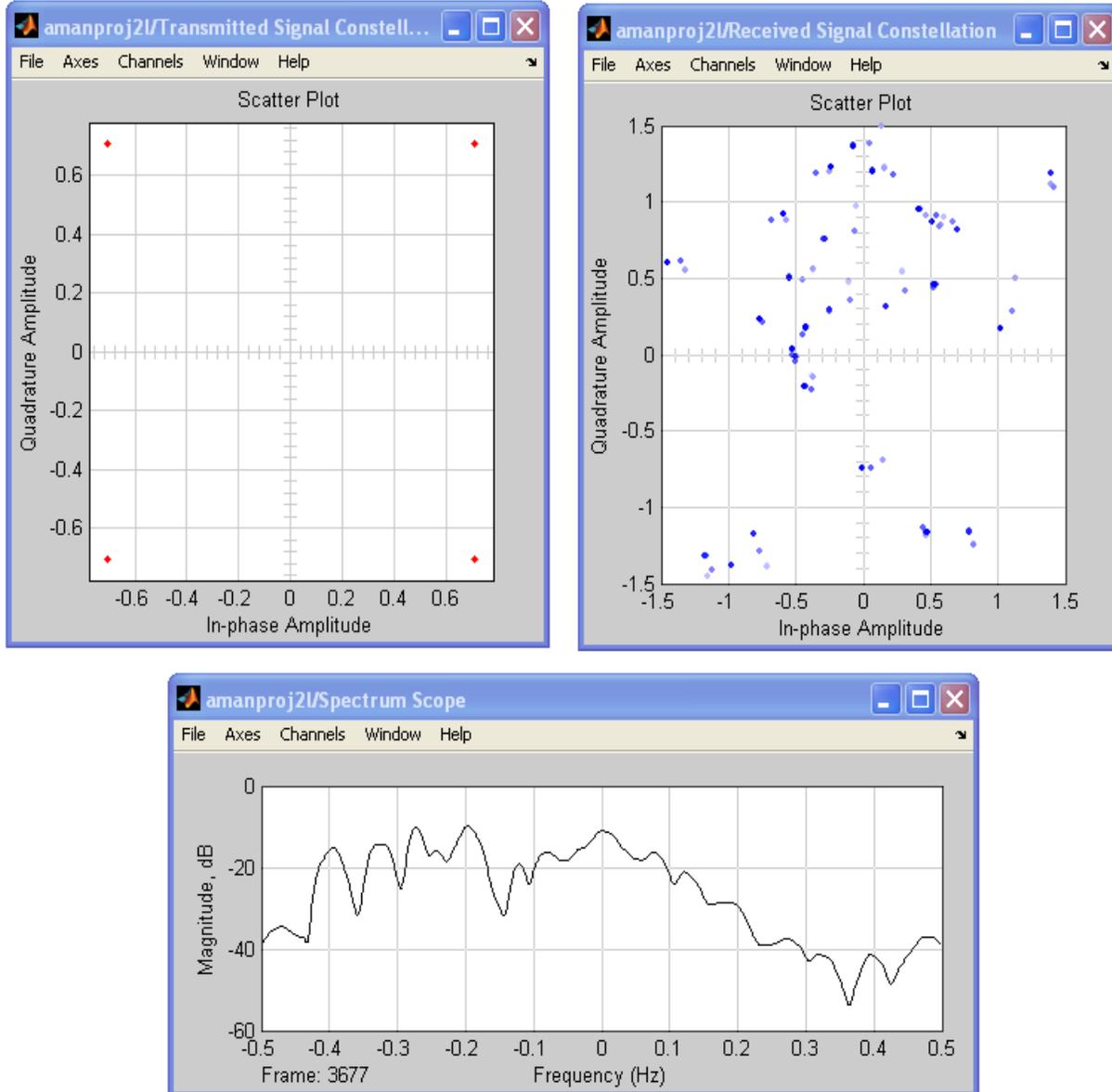
6. CONCLUSION

OFDM has already found wide deployment in various applications such as WLAN [3], vehicular communications [1] [2], etc. An OFDM link has been confirmed to work by using MATLABTM simulations, and some practical tests performed on a low bandwidth baseband signal. Four main performance criteria have been tested, which are OFDM's tolerance to multipath delay spread, channel noise, peak power clipping and start time error. Several other important factors affecting the performance of OFDM have only been partly measured. These include the effect of frequency stability errors on OFDM and impulse noise effects.

Most practical systems would use forward error correction to improve the system performance. Thus more work needs to be done on studying forward error correction schemes that would be suitable for telephony applications, and data transmission.
We can find many advantages in OFDM, but there are still many complex problems to solve.
This MATLABTM simulation proves that OFDM is better suited to a multipath channel than a single carrier

transmission technique such as 16-QAM

REFERENCES

[1] "Measurement Challenges for OFDM Systems", Agilent Technologies

[2] Denis Petrovic, Wolfgang Rave and Gerhard Fettweis, "Phase Noise Suppression in OFDM Using a Kalman Filter", Vodafone Chair for Mobile Communications, Dresden University of Technology, Germany

[3] http://digitalradiotech.co.uk/cofdm.htm [Digital Radio Tech]

[4] http://www.wi-fiplanet.com/tutorials/article.php/1500641

[5] http://www.ert.rwth-aachen.de/Projekte/Theo/OFDM/node1.html

[6] http://en.wikipedia.org/wiki/OFDM [Wikipedia]

[7] http://zone.ni.com/devzone/cda/tut/p/id/3740 [National Instruments]

[8] http://www.radio-electronics.com/info/cellulartelecomms/ofdm/ofdm.php

[9] Ove Edfors, Magnus Sandell, Jan Jaap Van de Beek, Daniel Landstorm and Frank Sjoberg, "An introduction to OFDM", 1996

[10] J.H.Scott, "Explaining some of the magic of COFDM", Proceedings of the 20th International Television Symposium, 1997

[11] "High Speed Wireless OFDM Communication Systems", Whitepaper, Wi-LAN Inc., 2001

[12] Tony Gomez, "OFDM for Mobile Data Communications", Web ProForum Tutorials, The International Engineering Consortium, 2001

[13] K. Panta and J. Armstrong, "Spectral Analysis of OFDM Signals and its Improvement by Polynomial Cancellation Coding", IEEE 2003

[14] Charan Langton, "Orthogonal Frequency Division Multiplexing", Intuitive guide to Principles of Communications, www.complextoreal.com, 2004

[15] Henrik Schulze and Christian Luders, "Theory and Applications of OFDM and CDMA", pp.145-264, 2005, John Wiley & Sons Ltd.

[16] Andreas F. Molisch, "Wireless Communcations", pp.399-424, 2005, John Wiley & Sons Ltd.

[17] L. Henzo and T. Keller, "OFDM and MC-CDMA: A Primer", pp.21-38, 2006, John Wiley & Sons Ltd.

SSDA.Analysis - A Class Library for Analysis of Sample Survey Data

Anu Sharma[1] and S. B. Lal[1]
[1](Indian Agricultural Statistics Research Institute) Pusa, New Delhi-110012

ABSTRACT

Agricultural researchers frequently use sample survey methodologies for estimation of various parameters in crops, livestock, fisheries production and allied fields. Analysis of data generated from these surveys requires the use of specialized software for survey data analysis. Most of the software for survey data analysis has proprietary source code and libraries which are not available to the users and cannot be utilized for the development of new applications. SSDA.Analysis is an objected oriented C# class library for analysis of survey data. The main goal of this library is to support fast development of MS-Windows based applications requiring readymade procedures for survey data analysis. It implements the logic of standard procedures for the estimation of parameters for various sampling designs within a framework designed to be easy to use, extend, and integrate with other .NET compatible software tools. This reusable library is highly useful for programmers and statistician involved in statistical software development. A windows based software named, Software for Survey Data Analysis (SSDA), has been developed using this library for the survey data analysis. We also reports here the results obtained after analyzing the mushroom data using this software.

Keywords: **Analysis, C#, library, sample survey.**

I. INTRODUCTION

Agricultural researchers frequently use sample survey methodologies for estimation of various parameters in crops, livestock, fisheries production and allied fields. Analysis of crop performance over a defined production area requires the identification and quantification of yield determining variables coupled with measurements of yield. Because of the inherent complexity of most crop production ecosystems, analyses of crop performance usually necessitate the collection and interpretation of large amounts of data describing yield and yield constraints and proponents. Analysis of this data requires the use of specialized software for survey data analysis.

Most of the statistical packages available worldwide either have more extensive features or are expensive. Softwares used for survey data analysis are SUDAAN, STATA, WesVarPC, PC-CARP, CENVAR and CLUSTERS etc. (Lepkowski J. and Bowles J. 1996). Some of these packages are windows based and others are DOS-based. SUDAAN is a commercial statistical software package for the analysis of correlated data, including correlated data encountered in complex sample surveys. SUDAAN is a single program consisting of a family of eleven analytic procedures used to analyze data from complex sample surveys and other observational and experimental studies involving repeated measures and cluster-correlated data (Wikepedia, 2008).

Most of the software has proprietary source code and libraries which are not available to the users. So their procedures can not be utilized in developing new software that requires similar types of computations. Also, this makes it impossible for users to extend functionality to these packages and prohibits the user for experimentation with or customization of sampling algorithms. At present no such ready made library and associated source code seems to be available for the analysis of sample survey data. So, there is a strong need of development of reusable library that implements the standard procedures for the estimation of parameters for various sampling schemes.

A set of C# library classes for, survey data analysis based software development, seems to be the proper solution to these problems, exploiting the modularity, reusability and versatility of C# design and coding.

The objective of this work was to investigate the use of object-oriented programming (OOP) as a tool for creation and management of survey data analysis procedures for agricultural applications. Specifically, we sought a programming approach that would allow (i) incremental application building without rewriting existing code, (ii) construction of a user-friendly interface from which all parameters can be assigned and software runs.

In this paper we presents SSDA.Analysis, an object oriented library for estimation of parameters of parameters of interest for Stratified Multistage Sampling Design. A windows based software and a sample web application have been developed using this library for the analysis of survey data.

II. ABOUT SSDA LIBRARY

SSDA.Analysis has been developed using C#.NET that uses .NET framework 2.0. C# is an object-oriented programming language and utilizes various key features of object oriented technologies such as its ability to program in an event driven operating system with great ease, write

code for events automatically, optimize code capability for native platform etc (Haertle, R. 2002; Robinson, S. et al. 2004). C# programming language was selected as it provides features that reduce common programming errors and is supported by a wealth of standard libraries and programming tools. SSDA.Analysis provides the methods for calculating the estimates of population mean, variance and design efficiency of the sampling scheme in comparison to the simple random sampling without replacement for the sampling designs shown in Table 1. The standard procedures as given in (Sukhatme, P.V et al. 1984) have been followed for estimation of population parameters for various sampling schemes.

SSDA.Analysis also provides methods for imputation of missing data using zero substitution, mean and mean of neighboring units. It also contains methods to provide descriptive statistics of the sampled data. Measures of central tendency included are arithmetic mean, median and measures of dispersion included are variance, co-efficient of variation, skewness and kurtosis.

III. DESIGN OF SSDA.ANALYSIS LIBRARY
SSDA.Analysis library is divided into following group of classes:

3.1 Simple Random Sampling (SRS)
Two classes namely SRS_EQ and ratio_random are available for SRS. SRS_EQ class implements the procedures for calculating mean, variance, relative standard error with/without replacement along with common methods like calculating total sum, sum of squares, total number of observations etc. in case of simple random sampling with equal/unequal probabilities. ratio_random class implements the procedure for estimating mean, bias, mean squared error with/without replacement with equal probabilities for the cases where auxiliary variable is available.

3.2 Stratified Sampling:
Three classes namely STRATIFIED, Stratified_UEQ_PR and ratio_random are available for stratified sampling design. STRATIFIED and Stratified_UEQ_PR classes implements the procedures for calculating mean, variance, relative standard error with/without replacement along with common methods like calculating total sum, sum of squares, total number of observations, number of stratum, total sum for each stratum etc. in case of stratified sampling with equal and unequal probabilities respectively. ratio_random class also contains the procedures for calculating mean, bias, mean squared error with/without replacement for stratified sampling design with equal probabilities for the cases where auxiliary variable is available.

3.3 Systematic Sampling
SYS_RANDOM implements the procedures for calculating mean, variance, relative standard error with/with replacement along with common methods like calculating total sum, sum of squares, total number of observations etc.

3.4 Cluster Sampling
Cluster class implements the procedures for calculating mean, variance, relative standard error with/with replacement along with common methods like calculating total sum, sum of squares, total number of observations, number of clusters, total sum for each cluster etc. in case of cluster sampling with equal/unequal probabilities.

3.5 Two Stage Sampling
two_stage class implements the procedures for calculating mean, variance, relative standard error with/without replacement with equal/ unequal probabilities for two stage sampling scheme.

3.6 Stratified Two Stage Sampling
STR_TWO_STAGE class implements the procedures for calculating mean, variance and relative standard error for stratified two stage scheme for without replacement with equal probabilities and for with replacement with unequal probabilities. Fig. 1 shows the class diagram of the SSDA.Analysis library.
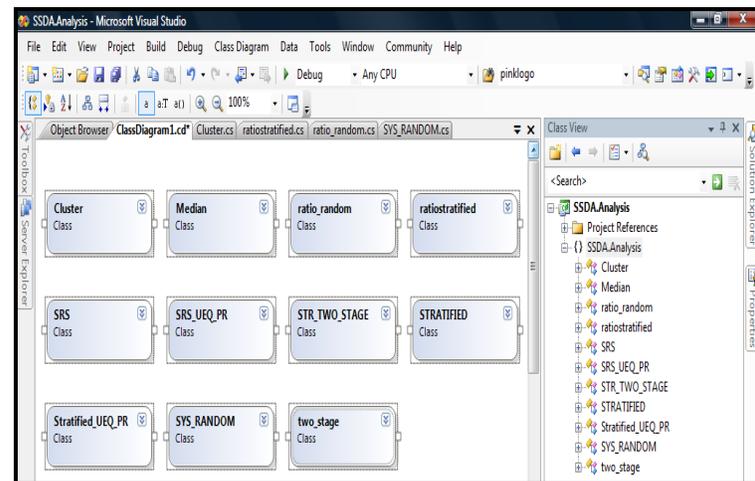

Fig.1: Class Diagram of SSDA.Analysis Library

IV. USING SSDA.ANALYSIS LIBRARY
Steps for using this library in .NET Applications are:

i.   Create a Console/Window/Web Application using Visual Studio.NET

ii. Click on Add Reference

iii. Select the SSDA.Analysis library. After adding reference to SSDA.Analysis library, it appears in namespace references.

iv. Use namespace by Adding using ssda.Analysis in the beginning for your project.

v. using ssda.analysis;

vi. Create an Object of ssda.

vii. ssda.Analysis cls = new ssda.Analysis ();

viii. Call Methods and Properties

V. APPLICATIONS OF SSDA.ANALYSIS

5.1 Windows based application development

A windows based Software for Analysis of Survey Data (SSDA) has been developed using the set of classes under SSDA.Analysis library. SSDA estimates the parameters for stratified multistage sampling design. It has four modules namely data management, analysis, report and HTML help. It has been developed using C# (C-Sharp) programming language available under .NET programming environment. This software is available on request at IASRI (www.iasri.res.in).

5.2 Web based application development

This reusable library may be utilized for the development of web based application and web services for analysis of survey data. Implementation of these functions allows separation of code from client applications. So, this allows us to develop many types of client applications like window or web based or web services.

VI. A CASE STUDY - USING SSDA SOFTWARE FOR ANALYZING MUSHROOM DATA OBTAINED FROM A REAL LIFE SURVEY

For illustration, we have considered the data obtained from a real life survey to estimate mushroom production (Gupta, Sud and Mathur 2009). The primary data have been collected in Sonepat district of Haryana state pertaining to Button Mushroom (Agaricus Bisporus) crop from November 2007 to April 2008. The data pertaining to shed area for raising mushroom, number of beds in each shed, weight of wet compost used, production of mushroom (q/ha), spawn consumed, wheat/paddy straw used in preparation of compost, processing of mushroom after picking, disposal of produce etc. have been collected from the selected mushroom growers in each of the selected village by enquiry method.

Stratified two-stage random sampling design with blocks/group of blocks as strata, mushroom–growing villages as primary stage sampling units and mushroom growing cultivar as the ultimate stage unit of selection. All the 6 blocks in the district were stratified into 3 strata by suitably combining the adjoining blocks. The three strata thus considered were Ganaur, Sonepat and Rai. A total of 8 villages, 3 from each of Ganaur and Sonepat and 2 from Rai were selected by simple random sampling without replacement. All the mushroom growers are in each of the selected villages were categorized into three categories as small, medium and large. Six cultivars were selected from these categories. Details are given in Table 1.

TABLE 1: Stratum-wise total number of mushroom growing villages, number and name of selected villages, total number of mushroom growers in the selected village (category-wise) and number of selected growers in each of the selected villages

Stratum	Total Number of Villages	Selected Villages	Name of the selected villages	Total number of mushroom growers	Selected mushroom growers
Ganaur (stratum – 1) (Ganaur & Gohana)	23	3	Ahirmajra	24 (5,9, 10)	6 (1,2,3)
			Ganaur	08 (7, 0,1)	6 (5, 0, 1)
			Rajlugarhi	23 (23, 0, 0)	6 (6, 0, 0)
Sonepat (stratum – 2) (Sonepat & Kharkhoda)	22	3	Rohat	26 (23, 0, 3)	6 (4, 0, 2)
			Kakroi	18 (18, 0, 0)	6 (6, 0, 0)
			Baiyapur	21 (6, 6, 9)	6 (2, 2, 2)
Rai (stratum – 3) (Rai & Mudlana)	08	2	Sersa	06 (0, 0, 6)	6 (0, 0, 6)
			Aterna	27 (27, 0, 0)	6 (6, 0, 0)
Total	53	08		153 (109,15, 26)	48 (30, 4,14)

The data collected is entered in MS-Excel file. This data is imported into SSDA software using file→Import option under data management module as shown Fig. 2. Stratified two stage sampling without replacement and with equal probability of selection was selected from various options provided in "select sampling design" dialogue box and input data was entered as shown in Fig. 3.

FIG. 2 Button Mushroom (Agaricus Bisporus) crop data in Sonepat district of Haryana from November 2007 to April 2008

FIG. 3 Screen for Entering Input Data

Results of analysis obtained using software for the variable production of mushroom (q/ha) are shown in Fig. 4. These results are tested with the results obtained after analyzing the same data using MS-Excel using same estimators. Both the results were found same. This ensures the reliability of the software for survey data analysis.

Population Sizes of PSUs

Stratum Number	Population Sizes
1	22
2	23
3	8

Estimate of Parameters

Study Variable	Stratum	Mean	Variance	RSE (%)*	Variance (SRS)	Design Efficiency
var_4	1	324.5448984	1,561.7499707	12.1767381	NA	NA
var_4	2	303.7120732	5,193.9962225	23.7295093	NA	NA
var_4	3	369.6038587	7,042.6840844	52.0732552	NA	NA
var_4	Pooled	321.7916873	1,812.1793432	13.2289619	70.1453405	3.8707725

FIG. 4: Results of Analysis

VII.   CONCLUSIONS

Analysis of survey data is an important part of all the surveys. Many software are available for analyzing the data obtained from agricultural surveys. Most of the software has proprietary source code and libraries which are not available to the users. So their procedures can not be utilized in developing new software that requires similar types of computations. Reusable class libraries developed here can be utilized for developing many other types of applications for analyzing survey data like web applications, web services. This library would be useful to computer scientists involved in statistical software development.

REFERENCES

[1] Gupta A.K., Sud U. C. and Mathur D.C. 2009, *Pilot Study to develop sampling methodology for estimation of production of mushroom crop*. Project Report, IASRI publication.

[2] R. Haertle, *OOP with Microsoft Visual Basic .NET and Microsoft Visual C# Step by Step* (Microsoft Press, 2002)

[3] Lepkowski J. and Bowles J. 1996, Sampling Error Software for Personal Computers. *The Survey Statistician*, 35, 10-17.

[4] Robinson, S., Nagel, C., Glynn, J., Skinner, M., Watson, K. and Evjen, B. 2004, *Professional C#*. 3rd Edition, Wiley Publishing.

[5] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S.and Asok, C., *Sampling Theory of Surveys with Application.* (Iowa State Univ. Press, Ames, Iowa and Indian   Society of Agricultural Statistics, New Delhi, 1984).

[6] Wikipedia(2008),   SUDAAN,   available   at http://en.wikipedia.org/wiki/SUDAAN.

Effect of Co-pesticide on Adsorption- Desorption Process on Agricultural Soils

Rounak M. Shariff

Department of Chemistry, College of Science / University of Salahaddine-Erbile, Kurdistan Region, Iraq

ABSTRACT

This work aim is to study the effect of different co-pesticides as Atrazine and propanil on adsorption processes on agricultural soil samples. The co-pesticides as Atrazine on adsorption behavior of metolachlor [2-chloro-N-   (2-ethyl-6-methylphenyl)-N-(2-methoxy-1-methylethyl) acetamide)] which is nonionic herbicide, and propanil on adsorption behavior of 2,4-D (2,4-dichlorophenoxyacetic acid) as anionic herbicide have been preformed, using batch equilibrium experiments on six agricultural soil samples. Linear and Freundlich models were used to describe the competitive sorption between the pair herbicides. Variation in adsorption affinities of the soils to the pesticides was observed. Freundlich coefficient K_F values for adsorption process varied between 0.079 - 2.282 mlg^{-1} and 0.058- 0.720 mlg^{-1} for metolachlor/atrazine and 2, 4-D/propanil respectively. The sorption strength of the herbicides decreased with increasing solution concentration. A nonionic surfactant was tested for its desorption potential and was found to be fairly effective at critical micelles concentration cmc concentration with removal of more than 65% sorbed pair competitive herbicides. Freundlich coefficient for desorption process K_{Fdes} for desorption process varied between 0.209- 0.523 and 0.926- 1.296 mlg^{-1} for metolachlor/atrazine and 2,4-D/propanil respectively. To investigate the effect of adsorption-desorption in the presence of cmc concentration of the nonionic surfactant for each pesticide alone also performed by using batch equilibrium experiments on six agricultural soil samples. The Freundlich coefficient K_F in the presence of the nonionic surfactant for metolachlor and 2,4-D ranged between 0.337-0.437 and 0.001-1.012 mlg^{-1} for adsorption processes. The Freundlich coefficient K_{Fdes} values for desorption process in presence of the nonionic surfactant ranged from 0.209 to 0.689 mlg^{-1} and 0.238 to 1.442 mlg^{-1} for metolachlor and 2,4-D respectively.

Keywords - Atrazine, 2, 4-D, HPLC, Metalachlor Propanil.

I. INTRODUCTION

Co-application of herbicides to soil created competitive sorption between the two herbicide producing smaller partition coefficients than for separate each herbicide. The herbicides co-application on agricultural fields has the potential to increase the mobility of these herbicides in soil, thereby also increasing the risk for groundwater concentration [1&2]. Contamination of these compounds with soil and drinking water has been generally recognized as dangerous [3]. Two major factors known to influence sorption of pesticides are soil properties and molecular characteristics. Sorption of neutral organic pesticides, as atrazine depends primarily on soil organic carbon (OC) content [4]. Sorption determines whether the pesticide will persist, be transported, and pollute the underlying ground water [5].   Strongly adsorbed and persistent pesticides that have large (K_{oc}) values are likely to remain near the soil surface. In contrast, weakly adsorbed but persistent pesticides (small K_{oc}) may be readily leached through the soil and more likely to contaminate ground water [6]. Kinetic data, which are measured infrequently, have the advantage of taking into account possible time-dependent reactions for adsorption, release, or desorption. Non-equilibrium conditions may be caused by the heterogeneity of sorption sites and slow diffusion to sites within the soil matrix, or organic matter [7]. Batch equilibrium experiments and Freundlich values are usually obtained for such competes for sorption sites in soil [8]. The selected herbicides co-application atrazine [9&10] and metolachlor [11] to soil. Each herbicides were persistence in the soil, its soluble in water, and poorly bound to most soils so it leaching down towards the ground water. The two nonionic herbicides were essentially slightly decreased and desorption amount increase little. Propanil [10&12] and 2, 4-D [11] herbicides co-application to soil. The effects of propanil on the sorption of anionic herbicides are possibly caused by the enhancement of electrostatic repulsion by pre-sorbed anionic herbicide/propanil and competition for sorption on interior sorption sites of soil particles which probably leads to small smaller partition coefficients. Employing of surfactant as batch washing techniques is an ex-situ process in which the contaminated soil is first excavated [13], which cause higher desorption for the herbicides.

II. MATERIALS AND METHODS
2.1 SOLIS

Fresh soil samples were taken from six soil samples were collected from six main agricultural, representing a range of physico-chemical properties. Subsamples of homogenized soils were analyzed for moisture content, organic matter content, particle size distribution, texture, pH, loss on ignition and exchangeable basic cations the detail were characterized in previous article[14].

2.2 MATERIALS

Analytical grad substituted with following purities expressed in weight percent Metolachlor (purity 97.8%), and 2, 4-D (purity 98%) atrazine (purity 99.2%), propanil (purity 99.7%) respectively. Were all purchased from Riedal-de Haen, Sigma-Aldrich company ltd. A nonionic surfactant TritonX-100 (TX-100), its chemical name is [Octylphenol ethoxcylate] surfactant, its Empirical formula is $(C_8H_{17}C_6H_4O(CH_2CH_2O)_NH)$; where N=9.5, its molecular weight is 625 g mol^{-1}, and its critical micelles concentration cmc concentration 0.0002M was obtained from Fluka AG, Buchs SG, and were used without further treatments. All chemicals used were of analytical grade reagents and used without pre-treatments. Standard stock solutions of the pesticides were prepared in deionised water.

2.3 ADSORPTION-DESORPTION EXPERIMENTS

Adsorption of the pesticides from aqueous solution was determined at temperature (25±1 C°) employing a standard batch equilibrium method. Duplicate air-dried soil samples were equilibrated with different pesticide concentrations (1, 2, 4 and 8 μg ml^{-1}) were for metolachlor/Atrazine 2,4-D/propanil, which the ratio of each pair were 1:1. The samples plus blanks (no pesticide) and control (no soil) were thermostated and placed in shaker for 0.5, 1, 3, 6, 9, 12 and 24h for metolachlor/Atrazine and 0.5, 1, 1.5,2, 2.5, 3, 3.5, 4, 6 and 24h for 2,4-D/propanil. The tubes were centrifuged for 20 min. at 3000 rpm. One ml of the clear supernatant was removed and analyzed for the pesticide concentration. Pesticide identification was done by PerkinElmer series 200 USA family high performance liquid chromatography (HPLC) for each pesticide concentration. The detailed information about the soil characteristics and their sorption process has been reported in previous work [11]. Desorption processes were done as each test tube was placed in a thermostated shaker at (25±1 C°) after equilibration for sufficient time as mentioned above with different pesticide concentrations (1, 2, 4 and 8 μg ml^{-1}) the samples were centrifuged, 5ml of supernatant was removed from the adsorption equilibrium solution and immediately replaced by 5ml of a nonionic surfactant at cmc concentration and this repeated for four times [15]. Adsorption-desorption of two pesticides each alone done in presence of nonionic surfactant at cmc concentration as mentioned above [16]. The resuspended samples were shaken for mentioned time previously for the kinetic study for each pesticide.

III. DATA ANALYSIS
3.1 ADSRPTION-DESORPTION ISOTHERM
3. 2.1 DISTRIBUTION COEFFICIENT

The distribution coefficient (K_d) was calculated by the using the following expression [17].

$$C_s = K_d C_e \qquad (1)$$

The distribution coefficient (K_d) was calculated by taking the ratio of adsorption concentration in soil (C_s) and equilibrium concentration in solution (C_e), and averaged across all equilibrium concentration to obtain a single estimate of K_d of the pesticides demonstrated in (Table 1-4)

3. 2.1 FREUNDLICH COEFFICIENT

Adsorption isotherm parameters were calculated using the linearized form of Freundlich equation [17]

$$\log C_s = \log K_F + \frac{1}{n}\log C_e \qquad (2)$$

C_s and C_e were defined previously, K_F is Freundlich adsorption coefficients, and n is a linearity factor, it is also known as adsorption intensity, 1/n is the slope and logK_F is the intercept of the straight line resulting from the plot of logC_s versus logC_e as shown in fig 1-4 . The values of K_F and 1/n calculated from this regression equation showed that Freundlich adsorption model effectively describes isotherms for the pesticides in all cases. Desorption isotherms of the pesticides were fitted to the linearzed form of the Freundlich equation [18].

$$\log C_s = \log K_{Fdes} + \frac{1}{n_{des}}\log C_e \qquad (3)$$

Where C_s is the amount of pesticides still adsorbed (μg g^{-1}), C_e is the equilibrium concentration of pesticides in solution after desorption (μg mL^{-1}), and K_{Fdes} (μg g$^{1-nfdes}$ /ml^{nfdes} g^{-1}) and n_{fdes} are two characteristic constants of the pesticides desorption [19]. The value of the K_{Fdes} and n_{fdes} constants of the pesticides demonstrated in (Table1-4).

3. 3 HYSTERESIS COEFFICIENT

A study of the pesticides desorption isotherms show positive hysteresis coefficients H_1 on the six selected soil samples. Hysteresis coefficients (H_1) can be determined by using the following equation [18].

$$H_1 = \frac{n_a}{n_{des}} \qquad (4)$$

Where n_a and n_{des} are Ferundlich adsorption and desorption constants, respectively, indicating the greater or lesser irreversibility of adsorption in all samples, the highest values corresponding for which the highest adsorption constant was obtained. The coefficient H_1 is a simple one and easy to use, Data in table 5&6 demonstrated H_1 values for metolachlor, and 2, 4-D respectively. The extent of hysteresis was quantified by using hysteresis coefficient (ω), it was defined on the discrepancy between the sorption and desorption isotherms, and calculated by using Freundlich parameters estimated from sorption and desorption isotherms separately, (ω) expressed as [20].

$$\omega = (\frac{n_a}{n_{des}} - 1)x100 \qquad (5)$$

Recently Zhu et. al [21] proposed an alternative hysteresis coefficient (λ) based on the difference in the areas between adsorption and desorption isotherms, they derived the

following expression for the parameter λ for the traditional isotherms.

$$\lambda = (\frac{n_a + 1}{n_{des} + 1} - 1) x 100 \qquad (6)$$

3. 4 ORGANIC MATTER NORMALIZED ADSORPTION COEFFICIENT

The linear or distribution coefficient (K_d) is related to soil organic carbon (OC) and soil organic matter (OM) by the following equations [22].

$$\%OC = \frac{\%OM}{1.724} \qquad (7)$$

$$K_{OM} = \frac{100K_d}{\%OM} \qquad (8)$$

$$K_{OC} = \frac{100K_d}{\%OC} \qquad (9)$$

IV. RESULTS AND DISCUSSION

The K_d values for adsorption process for metolachlor/atrazine varied between 1.384 - 2.832 mlg^{-1} while the regression coefficient R^2 value ranging from 0.705 to 0.895 with standard error S.E. value between 0.012 – 0.096. The K_d values for adsorption process for 2,4-D/propanil varied between 2.209- 5.542 mlg^{-1}, while the value of R^2 ranging from 0.704 to 0.783 with standard error S.E. value between 0.019 – 0.058, the regression equations relating that the highest values are the most fitted model. The competitive effect of pair pesticides sorption by soil generally decreased with increasing initial herbicide concentrations because of the saturation of sorption sites in soil, our results agreed with research [2].The each two pair pesticides are widely used for pest control in agricultural crops[23&24].

The desorption experiments were conducted with a nonionic surfactant TritonX-100 at concentration 0.1cmc, cmc and 20cmc on metolachlor and 2,4-D sorbed soil corresponding to initial concentration 4 µg mL^{-1} the comparative results shown in table 7, present that the degree of desorption of each pesticides from soil into surfactant solution was 46 % for 0.1cmc concentration; 65% for cmc concentration and 68% for 20cmc concentration. The cmc concentration gave the best results as there was only a normal increase in desorption at the concentration of 3%. So the used surfactant solution is therefore fairly effective in desorption of metolachlor and 2, 4-D from the contaminated soil. The dynamics are believed to be highly related to the adsorption of the four herbicides used, could be partially explained by the unoccupied sites in each soil more by the total sorption capacity [25]. Although results of the research showing an

important role to the organic matter in the adsorption process[26]. Another explanation of the mechanism for adsorption process was the formation of adducts between the pair herbicide used and the and the constituent of the soil [27].

The K_d values for desorption process in the presence of cmc concentration of the surfactant varied between for metolachlor/atrazine 3.498- 14.43 mlg^{-1} while the value of R^2 ranging from 0.726 to 0.957 with standard error S.E. value between 0.016 – 0.090. The K_d values for desorption process in the presence of cmc concentration of the surfactant for 2,4-D/propanil varied between 18.25- 52.46 mlg^{-1} while the value of R^2 ranging from 0.763 to 0.996 with standard error S.E. value between 0.005 – 0.091. The Ferundlich nonlinear sorption isotherm showed a good fit to the measured data for all soil samples for metolachlor/atrazine and 2,4-D/propanil. The K_F values for adsorption process for metolachlor/atrazine varied between 0.079 - 2.282 mlg^{-1} , the R^2 value ranging from 0.758 to 0.889 with S.E. 0.032-0.039 and the value of the nonlinearity reneging between n_F 1.515-2.392. The K_F values for adsorption process for2,4-D/propanil 0.058- 0.720 mlg^{-1}, the R^2 value ranging from 0.799 to 0.987 with S.E. 0.032-0.038 and the value of the nonlinearity reneging between n_F 0.543-2.198. The values of K_F less for the two pair as the two pesticides were used alone. Freundlich coefficient for desorption process K_{Fdes} for metolachlor/atrazine in the presence of cmc concentration of the surfactant varied between 0.209- 0.523 mlg^{-1} the R^2 value ranging from 0.770 to 0.941 with S.E. 0.044-0.054, the values of n_{Fdes} 1.209-3.968. The K_{Fdes} for 2, 4-D/propanil in the presence of cmc concentration of the surfactant between 0.926- 1.296 mlg^{-1} the R^2 value ranging from 0.725 to 0.996 with S.E. 0.068-0.083, values of n_{Fdes} 0.588-1.342. The K_d values for adsorption process in the presence of cmc concentration of the surfactant varied between for metolachlor 1.000- 1.226 mlg^{-1} while the value of R^2 ranging from 0.703 to 0.901 with standard error S.E. value between 0.154 – 0.252. The K_d values for adsorption process in the presence of cmc concentration of the surfactant for 2,4-D varied between 1.332- 3.712 mlg^{-1} while the value of R^2 ranging from 0.693 to 0.965 with standard error S.E. value between 0.025 – 0.068. The K_d values for desorption process in the presence of cmc concentration of the surfactant varied between for metolachlor 4.431- 13.01 mlg^{-1} while the value of R^2 ranging from 0.768 to 0.957 with standard error S.E. value between 0.010 – 0.031. The K_d values for desorption process in the presence of cmc concentration of the surfactant for 2,4-D varied between 1.081- 57.48 mlg^{-1} while the value of R^2 ranging from 0.707 to 0.965 with standard error S.E. value between 0.023 – 0.086. The K_F values for adsorption process in the presence of cmc concentration of the surfactant for metolachlor varied between 0.337 - 0.437 mlg^{-1} , the R^2 value ranging from 0.752 to 0.880 with S.E. 0.041-0.049 and the value of the nonlinearity reneging between n_F 1.600-2.387. The K_F values for adsorption process in the presence of cmc concentration of the surfactant for 2, 4-D 0.001- 1.012 mlg^{-1}, the R^2 value ranging from 0.849 to 0.992 with S.E. 0.0324-0.039 and the value of the nonlinearity reneging between n_F 0.808-2.028.

Freundlich coefficient for desorption process K_{Fdes} for metolachlor in the presence of cmc concentration of the surfactant varied between 0.209- 0.689 mlg^{-1} the R^2 value ranging from 0.776 to 0.998 with S.E. 0.027-0.063, the values of n_{Fdes} 0.709-1.828. The K_{Fdes} for 2, 4-D in the presence of cmc concentration of the surfactant 0.238- 1.442 mlg^{-1} the R^2 value ranging from 0.735 to 0.976 with S.E. 0.021-0.089 ,the values of n_{Fdes} 0.888-1.486. The size of the organic is considered to play an important role in its rate of adsorption, Baily and white 1970[28]. et. al. summarized the of molecular size as follows: a)Adsorption of noneletrolytes by nonpolar adsorbents increases as molecular weights of the substances increases. b) Van der Waals forces of adsorption increases with increasing molecular size. c) Adsorption decreases because of steric hindrance. The evidence available also shows the presence of a maximum limit in molecular size in adsorption of organic compounds. Larger molecules (chain length greater than five units) may be adsorbed only in the presence of excess water. However very large molecules difficulties in adsorption due to adverse molecular configuration. The use of mixture for the each pair herbicides, This indicated that atrazine and propanil competed with the two herbicides for sorption sites in soil[29]. The differences in adsorption coefficient between samples were statistically significant, for soils with organic content for that readily sorbed propanil than the other, thereby being more competitive in soils with greater soil organic matter content. The Ferundlich slope of the isotherm was always less than unity and indicating that the affinity between the herbicides used alone and soil was greatest at initial herbicides concentrations and decreased as increasing the herbicides concentrations [30].

Data in table 5 demonstrated H_1 values for metolachlor/atrazine and 2,4-D/propanil from the selected soil samples in the range from 0.615-1.751 and 0.594-2.524 respectively. The calculated values of hysteresis coefficient (ω) for adsorption-desorption of for metolachlor/atrazine and 2, 4-D/propanil on the selected soil samples ranged from -43 to -75 and from -41 to 152 respectively. Whereas hysteresis coefficient (ω) is only applicable for the traditional type isotherms of the successive desorption [31&32]. The hysteresis coefficient (λ) for metolachlor/atrazine and 2,4-D/propanil from the selected soil samples were ranged from -10 to 148 and from 29 to 1662 respectively. The H_1 values for metolachlor and 2, 4-D alone in the presence of cmc surfactant as summarized in table 6 from the selected soil samples in the range from 1.144-2.687 and 0.869-2069 respectively, indicating an increase in the irreversibility of the adsorption of herbicide as the clay content increases, and indicate the increased difficulty of the sorbed analytic to desorbed from the matrix. The calculated values of hysteresis coefficient (ω) for adsorption-desorption of for metolachlor and 2,4-D on the selected soil samples ranged from 14 to 169 and from -13 to -107 respectively. The hysteresis coefficient (λ) for metolachlor and 2, 4-D from the selected soil samples were ranged from -40 to -94 and from -6 to 1441 respectively.

V. CONCLUSION

The using of each pair of herbicides metolachlor/atrazine and 2,4-D/propanil may increase herbicides leaching to depth relative to the use of each one alone on the six agricultural soil samples. The cmc concentration gave the best results in desorption. So the used surfactant solution is therefore fairly effective in desorption of metolachlor, 2, 4-D from the contaminated soil and for each pair.

REFERENCES

[1] Farenhorst A and Bowman B. T., Competitive sorption of atrazine and metolachlor in soil, *Journal of Science and health, 33, 1998,* 671-682.

[2] Annemieke Farenhorst and Brent Prokopowich. The effect of propanil Co-application on 2,4-D sorption by soil . Journal of environmental science and health part b-pesticides, *food contaminants, and agricultural wastes. 38(6), 2003,* 713-721.

[3] Pimentel, D., and L. Levitan, Amounts applied and amounts reaching pests, *Bioscience, 36, 1986, 86-91.*

[4] Madhun, Y. A., V. H. Freed, J. L., and Fang, Sorption of bromacilm chlortoluron, and diuron by soils, *Soil Sci. Soc. Am. J, 50, 1986,* 1467-1471.

[5] Rao, P. S. C. , A. G. Hornsby. 1983. Behavior of pesticides in soils and water. *Soil Sci. Fact Sheet SL40. Univ. of Florida, Gainesville.*

[6] E. Barriuso, Ch. Feller, R.Calvet and C.Cerri. Sorption of atrazine, terbutryn and 2,4-D herbicides in two Brazilian Oxisols. *Eleesevier Science Pulishers B. V.,Amsterdam, 53, 1992,* 155-167.

[7] P. Nkedi – kizza, D. Shinde, M. R.Savabi, Y. Ouyang, and L. Nieves. Sorption kinetics and equlibria of organic pesticides in cabonatic soils from south Florida. *J. Environ. Qual, 35, 2006,* 268-276.

[8] Hamaker J W., The Application of Mathematical Modeling to the Soil Persistence and Accumulation of Pesticides, Proc. BCPC Symposium, *Persistence of Inscticide and Herbicides, 1967,* 181-199.

[9] Benny Chefetz, Sorption of Phenantgrene and Atrazine by Plant Cuticular fractions, *Environmental Toxicology and Chemistry, 22, 2003,* 2492-2498.

[10] Kafia M. Shareef and Rounak M. Shariff, Adsorption of Herbicides on eight Agricultural Soils, *J. of university of anbar for pure science, 3, 2009*, 67-75.

[11] Rounak M. Shariff. "Adsorption-Desorption of Metolachlor and 2,4-D on Agricultural Soils". *International Journal of Scientific & Engineering Research, 2, 2011*, 1-8.

[12] Jack R. Plimmer, Philip C. Kearney, Hideo Chisaka, Joseph B. Yount, and Ute I. Klingebiel. 1,3- Bis (3,4-dichlorophenol) triazene from propanil in soils.J. AG, *FOOD CHEM.,18(5), 1970,*859-865.

[13] DiCesare, D., and J. A. Simth. 1994. Surfactant effects on desorption of nonionic compounds. *Rev. Environ. Contam. Toxicol.* 134:1-29.

[14] M. Rounak Shariff., Compost Adsorption Desorption of Picloram in the Presence of Surfactant on Six Agricultural Soils, *International Journal of Scientific & Engineering Research, 2(5), 2011,*2229-5518.

[15] Mohammed A Ali and Peter J. Baugh, Sorption, Desorption Studies of Six Pyerethroids and Mirex on Soils using GC/ MS-NICI , Internet. *J. Environ. Anal. Chem., 83(11),2003, 923-933.*

[16] E. Iglesias-Jimenez, M. J. Sanchez-Martin, M.Sanchez-Camazano, Pesticide adsorption in a soil – water system in the presence of surfactants, *Chemosphere, 32(9), 1996,* 1771-1782.

[17] OP Banasal., Kinetics of interaction of three carbamate pesticides with Indian soils: Aligrah district, *Pest Manag Sci., 60, 2004,* 1149-1155.

[18] R. A. Griffin and J. J. Jurinak, Test of a New Model for the Kinetics of Adsorption-Desorption Processes, *Soil Sci Soc. Amer. Proc., 37,1973,* 869-872.

[19] Zhu, H., and H. M. selim, Hysteretic of metolachlor Adsorption - desorption in soil*, J. Soil. Sci Qual.,165, 2000,* 632-645.

[20] Cass T. Miller and Joseph A. Pedlt, Sorption-Desorption Hysteresis and A biotic Degradation of Lindane in a Surface Material, *Environ. Sci. Technol, 26, 1992, 1417-1427.*

[21] H. M. Selim and H. Zhu, Organic Compounds in Environment Atrazine Sorption-Desorption Hysteriesis by Sugarcane Mulch Residue, *J. Environ. Qual., 34, 2005,* 325-335.

[22] Adlophe Monkiedje and Micheal Spiteller. fungicides, mefenoxam and meta laxly, and their acid metabolite in typical Cameroonian and German soils, *chemospher,49(6) 2002, 659-668.*

[23] T. Sismanoglu, A. Ercage, S. Pura and E. Ercage, Kinetics and isotherms of Dazomet Adsorption on Natural Adsorbents, *J. Braz. Chem. Soc., 15(5), 2004,* 669-675.

[24] Miels Henrik Spliid, Arne Helweggcnol and Kiesten Helnrichson "Leaching and degradation of 21 pesticides inafull-scale model biobed" 2006. Interscices.

[25] Thanh H. Dao, Competitive anion sorption effects on dairy wast water dissolved phosphorus extraction with zeolite-based sorbents, *Food, Agriculture & Environment,3(4), 200,3*263-269.

[26] Marcelo Kogan, Alejandia Metz and Rodrigo Ortega. , Adsorption of glyphosate in chilecan and its relationship with unoccupied phosphaste binding sites, *Pesq agropes bras .Brasili, 38(4), 2003,* 513-519.

[27] Baoshan Xing, Joseph J. Pingnatello, and Barbara Giglitti. Competitive sorption between atrazine and other organic compounds in soils and model sorbents . *Environ. Sci. Techno, 30, 1996,* 2432-2440

[28] Bailey, G. W. And J. L. White "Factors influencing the adsorption and movement of pesticides in soils "Springs New York .1970. vol 30 : 29-92.

[29] Daniel Said- Pullicino, Giovanni Gigliotti, and Alfred J. Vella, Environmental of triasulfuran in soils Amended with municipal waste compost, J. *Environ. Qual. ,33, 2004,* 1743-1751.

[30] Elsayed A. Elkhatib, A. M. Mahdy and N. H. Bbrakat, Thermodynamics of copper desorption from soils as affected by citrate and succinate, *Soil & water Res., 2 ,(4), 2007,*135-140.

[31] Cludio A. Spadotto and Arthur G. Hornsby, organic compounds in the environment soil sorption of acidic pesticides: modeling pH effects, *Environ Qual. 32, 2003,* 949-956.

Table 1: Adsorption of the co-application metolachlor/atrazine and their desorption in the presence of TritonX-100 at cmc concentration, the linear and Freundlich models isotherm parameters on the selected soil samples.

Ads-des Models	Parameter	Soils					
		S_1	S_2	S_3	S_4	S_5	S_6
Ads.Distr. coffi.	K_d (calc)	2.832	1.618	1.599	1.384	1.870	1.956
	S.E	0.075	0.072	0.096	0.012	0.075	0.095
	R^2	0.705	0.821	0.794	0.890	0.893	0.755
	K_{OC}(mL/g)	101	156	50	59	98	130
	K_{OM}(mL/g)	1.744	2.685	0.863	1.013	1.684	2.235
Freundlich (ads)	K_F(mL/g)	0.079	0.199	0.233	0.282	0.155	0.218
	S.E	0.034	0.037	0.038	0.039	0.032	0.035
	n_F	2.392	1.515	1.761	1.789	1.563	2.257
	R^2	0.864	0.835	0.810	0.889	0.776	0.758
Des.Distr. coffi.	K_d (calc)	4.158	8.386	3.498	14.43	6.116	4.366
	S.E	0.019	0.027	0.023	0.016	0.090	0.024
	R^2	0.940	0.780	0.957	0.780	0.759	0.726
Freundlich(des)	K_{Fdes}(mL/g)	0.244	0.461	0.209	0.464	0.523	0.252
	S.E	0.044	0.046	0.045	0.054	0.051	0.047
	n_F	1.366	2.463	1.178	2.463	1.209	3.968
	R^2	0.941	0.838	0.906	0.838	0.913	0.770

Table 2: Adsorption of metolachlor alone and it's desorption in the presence of TritonX-100 at cmc concentration, the linear and Freundlich models isotherm parameters on the selected soil samples.

Ads-des Models	Parameter	Soils					
		S_1	S_2	S_3	S_4	S_5	S_6
Ads.Distr. coffi.	K_d (calc)	1.142	1.160	1.099	1.000	1.226	1.014
	S.E	0.252	0.127	0.122	0.242	0.154	0.250
	R^2	0.741	0.800	0.901	0.703	0.777	0.721
	K_{OC}(mL/g)	40.8	112	34	42	64	67
	K_{OM}(mL/)	0.703	1.924	0.593	0.732	1.104	1.158
Freu	K_F(mL/g)	0.385	0.340	0.348	0.419	0.337	0.437

		S₁	S₂	S₃	S₄	S₅	S₆
	S.E	0.045	0.042	0.041	0.045	0.043	0.049
	n_F	2.092	1.718	1.600	2.387	1.905	2.325
	R^2	0.771	0.880	0.879	0.856	0.854	0.752
Des.Distr. coffi	K_d (calc)	11.11	13.01	4.431	7.109	8.069	12.17
	S.E	0.016	0.017	0.031	0.027	0.012	0.010
	R^2	0.868	0.768	0.859	0.781	0.957	0.811
Freundlich(des)	K_{Fdes}(mL/g)	0.658	0.659	0.209	0.488	0.614	0.689
	S.E	0.057	0.063	0.044	0.027	0.062	0.061
	n_F	1.828	0.987	1.178	2.049	0.709	1.451
	R^2	0.981	0.803	0.906	0.776	0.998	0.856

Table 3: Adsorption of the co-application of 2,4-D/propanil and their desorption in the presence of TritonX-100 at cmc concentration, the linear and Freundlich models isotherm parameters on the selected soil samples.

Ads-des Models	Parameter	Soils					
		S₁	S₂	S₃	S₄	S₅	S₆
Ads.Distr. coffi.	K_d (calc)	2.678	3.362	2.209	5.542	4.565	3.519
	S.E	0.058	0.019	0.064	0.029	0.035	0.033
	R^2	0.704	0.783	0.771	0.742	0.728	0.737
	K_{OC}(mL/g)	96	324	69	235	238	233
	K_{OM}(mL/g)	1.649	5.578	1.191	4.055	4.112	4.020
Freundlich (ads)	K_F(mL/g)	0.720	0.100	0.135	0.203	0.122	0.058
	S.E	0.034	0.035	0.033	0.038	0.036	0.037
	n_F	0.543	1.484	1.988	2.151	2.198	1.795
	R^2	0.799	0.987	0.920	0.872	0.935	0.806
Des.Distr. coffi.	K_d (calc)	18.25	30.15	27.56	28.04	52.46	21.56
	S.E	0.091	0.074	0.033	0.005	0.027	0.035
	R^2	0.835	0.838	0.912	0.996	0.763	0.828
Freundlich(des)	K_{Fdes}(mL/g)	0.926	1.296	1.096	1.131	1.291	1.022
	S.E	0.068	0.079	0.074	0.073	0.083	0.069
	n_F	0.914	0.588	1.114	1.043	1.342	0.968

R^2		0.782	0.932	0.891	0.996	0.725	0.897

Table 4: Adsorption of 2,4-D alone and its desorption in the presence of TritonX-100 at cmc concentration, the linear and Freundlich models isotherm parameters on the selected soil samples.

Ads-des Models	Parameter	Soils					
		S_1	S_2	S_3	S_4	S_5	S_6
Ads.Distr. coffi.	K_d (calc)	1.332	1.801	1.401	3.712	3.037	2.212
	S.E	0.068	0.025	0.067	0.042	0.041	0.033
	R^2	0.965	0.791	0.977	0.693	0.825	0.703
	K_{OC}(mL/g)	48	173	44	158	159	147
	K_{OM}(mL/g)	0.820	2.988	0.756	2.716	2.736	2.527
Freundlich (ads)	K_F(mL/g)	0.253	1.012	0.266	0.059	0.001	0.110
	S.E	0.038	0.036	0.037	0.039	0.035	0.034
	n_F	1.420	0.808	1.678	2.028	1.838	1.773
	R^2	0.930	0.977	0.992	0.905	0.849	0.922
Des.Distr. coffi.	K_d (calc)	1.081	21.27	13.54	16.77	57.48	27.954
	S.E	0.056	0.029	0.060	0.086	0.028	0.023
	R^2	0.965	0.932	0.888	0.707	0.838	0.802
Freundlich(des)	K_{Fdes}(mL/g)	0.238	1.031	0.754	0.784	1.442	1.064
	S.E	0.038	0.067	0.021	0.026	0.089	0.071
	n_{Fdes}	1.443	0.929	1.504	1.486	0.888	1.340
	R^2	0.976	0.927	0.999	0.735	0.796	0.932

Table 5: Hysteresis effect for the co-application of each pair metolachlor/atrazine, and 2,4-D/propanil on the selected soil samples.

Soil	Metolachlor /atrazine			2,4-D/propanil		
	H_1	ω	λ	H_1	ω	λ
S_1	1.751	75	184	0.594	-41	29
S_2	0.615	-39	109	2.524	152	1196
S_3	1.495	49	-10	1.785	79	712

S_4	0.726	-27	63	2.062	106	457
S_5	1.293	29	237	1.638	64	958
S_6	0.569	-43	16	1.854	85	1662

Table 6:Hysteresis effect for metolachlor and, 2,4-D each one alone on the selected soil samples.

Soil	metolachlor			2,4-D		
	H_1	ω	λ	H_1	ω	λ
S_1	1.144	14	71	0.984	-2	-6
S_2	1.740	74	94	0.869	-13	2
S_3	1.358	36	- 40	1.116	12	183
S_4	1.164	16	16	1.365	37	1228
S_5	2.687	169	82	2.069	107	1441
S_6	1.602	60	58	1.323	32	867

Table 7:Desorption data at initial sorbate concentration 4 µg ml^{-1} and pH=6 for metolachlor and, 2,4-D each one alone at different surfactant concentration.

Agitation Time (h)	Metolachlor at different Surfactant Concentration			2,4-D at different Surfactant Concentration		
	0.1cmc	cmc	20cmc	0.1cmc	cmc	20cmc
0.5	15.99	19.55	21.66	12.56	16.43	18.87
1	19.11	30.78	36.25	16.32	27.45	33.31
2	26.65	40.23	43.27	23.32	37.13	40.67
3	37.68	46.74	47.34	34.22	43.11	44.87
4	40.43	48.86	50.22	37.34	44.23	47.89
6	41.56	50.44	51.46	38.17	47.55	48.43
24	41.65	50.89	51.85	38.74	47.74	48.89
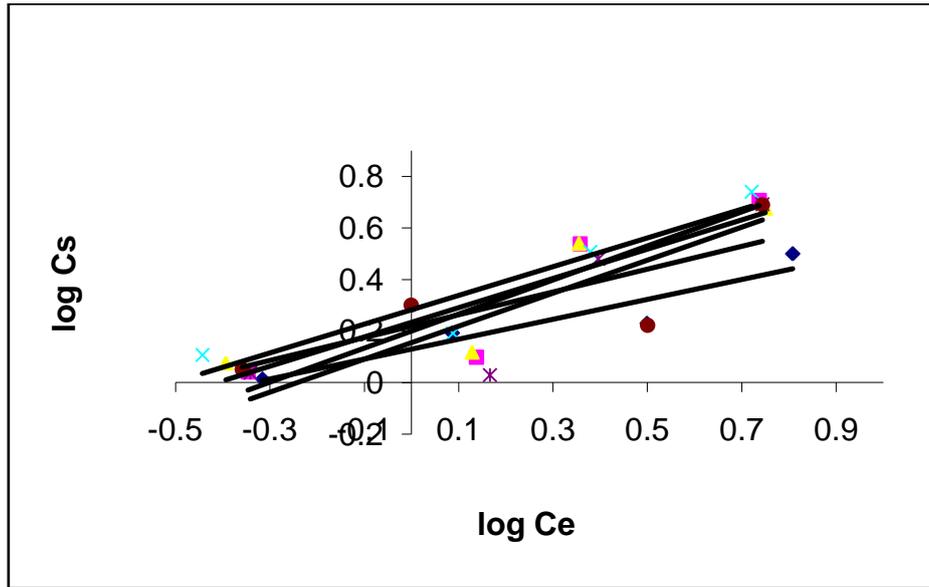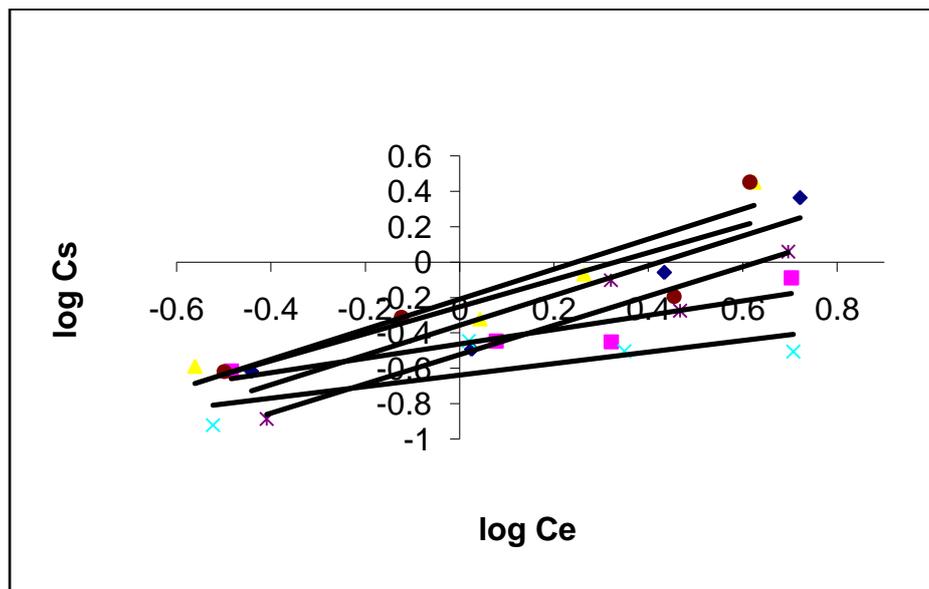
a-

b-

Fig. 1: Fitted Ferundlich model for co-application of metolachlor/atrazine (a) adsorption (b) desorption isotherm in the presence of nonionic surfactant on selected soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, ●S_6).
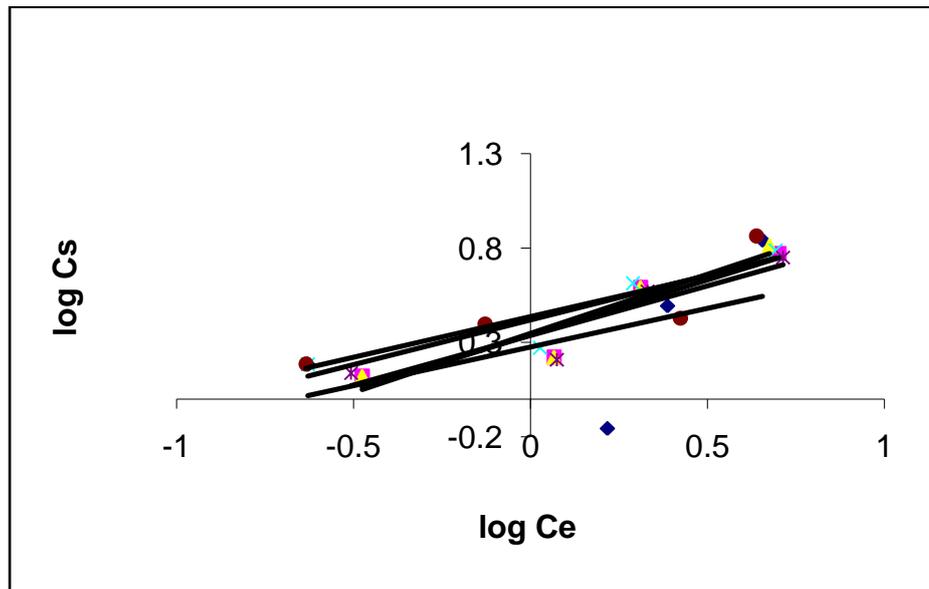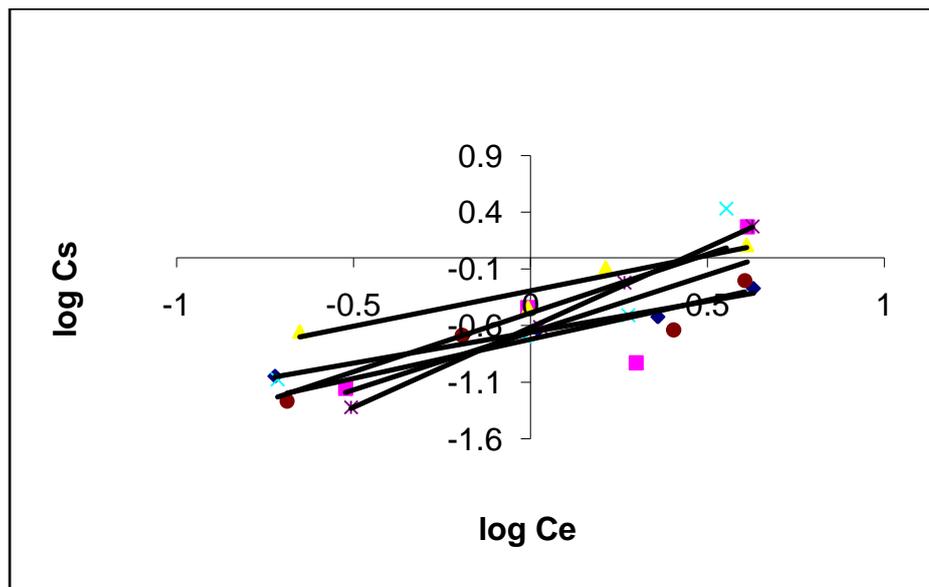
a-

b-

Fig. 2: Fitted Ferundlich model for metolachlor alone in presence of nonionic surfactant (a) adsorption (b) desorption isotherm selected soil samples (♦ S$_1$, ■ S$_2$, ▲ S$_3$,   x S$_4$, * S$_5$, ●S$_6$).
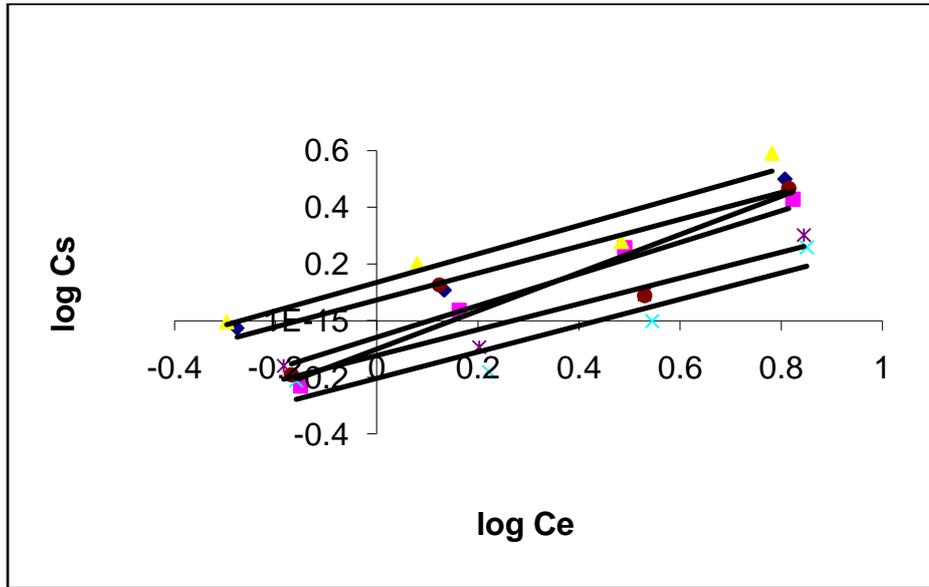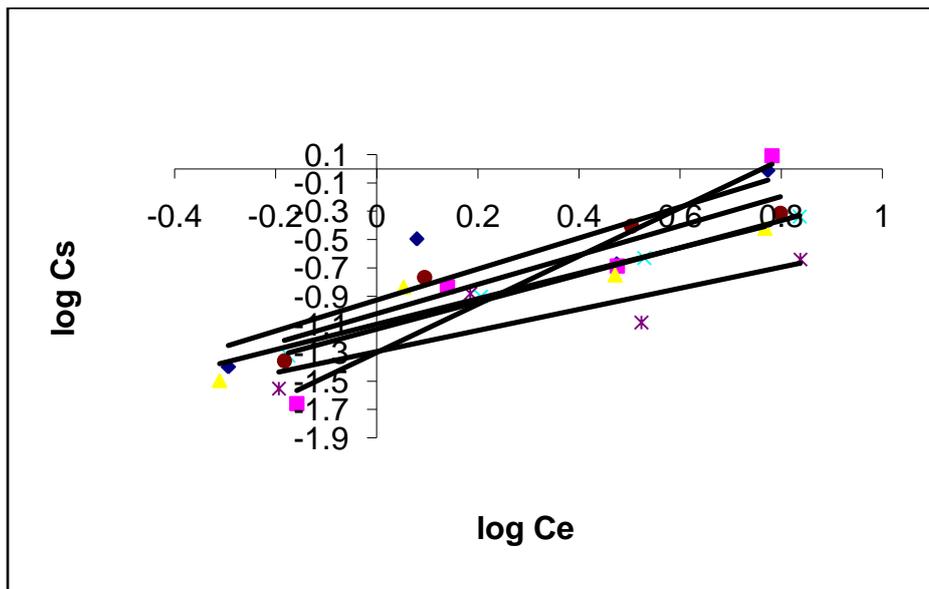
a-

b-

Fig. 3: Fitted Ferundlich model for co-application of 2,4-D/propanile (a) adsorption (b) desorption isotherm in the presence of nonionic surfactant on selected soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, ●S_6).
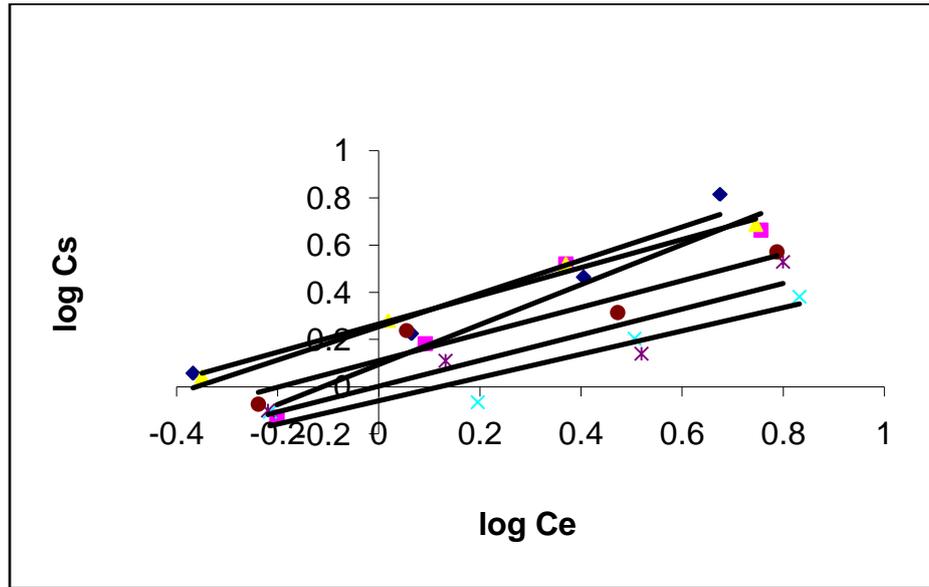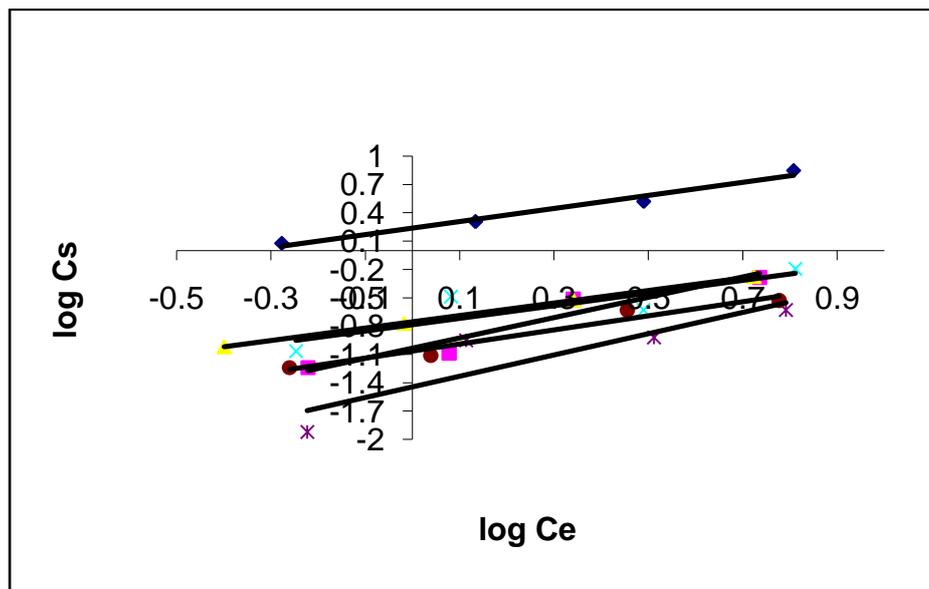
a-

b-

Fig. 4: Fitted Ferundlich model for 2,4-D alone in the presence of nonionic surfactant (a) adsorption (b) desorption isotherm on selected soil samples (♦ S_1, ■ S_2, ▲ S_3, x S_4, * S_5, ●S_6).

RANKING CONCEPT-BASED USER PROFILE FROM SEARCH ENGINE LOGS

R.Kokila[1] and Mrs.N.Krishnammal[2]

[1]M.E Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India
[2]Assistant Professor, Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India

ABSTRACT:

Commercial search engines return roughly the same results for the same query, regardless of the user's real interest. Since queries submitted to search engines tend to be short and ambiguous, they are not likely to be able to express the user's precise needs. In existing system, most existing user profiling strategies are based on objects that users are interested in (i.e., positive preferences), but not the objects that users dislike (i.e., negative preferences). Experimental results show that profiles which capture and utilize both of the user's positive and negative preferences perform the best and also negative preferences can increase the separation between similar and dissimilar queries. The separation can be achieved by using agglomerative clustering algorithm to terminate and improve the overall quality of the resulting query clusters. In the proposing system, queries submitted to search engines, they are likely to be able to express the user's precise needs and the concept based user profiles can be integrated into the ranking algorithms of a search engine so that search results can be ranked according to individual user's interests. This technique improves a search engine's performance by identifying the information needs for individual users.

Index terms- Negative preferences, personalization, personalized query clustering, search engine, user profile.

I.INTRODUCTION

MOST commercial search engines return roughly the same results for the same query, regardless of the user's real interest. Since queries submitted to search engines tend to be short and ambiguous, they are not likely to be able to express the user's precise needs. For example, a farmer may use the query "apple" to find information about growing delicious apples, while graphic designers may use the same query to find information about Apple Computer.

Personalized search is an important research area that aims to resolve the ambiguity of query terms. To increase the relevance of search results, personalized search engines create user profiles to capture the users' personal preferences and as such identify the actual goal of the input query. Since users are usually reluctant to explicitly provide their preferences due to the extra manual effort involved, recent research has focused on the automatic learning of user preferences from users' search histories or browsed documents and the development of personalized systems based on the learned user preferences.

A good user profiling strategy is an essential and fundamental component in search engine personalization. We studied various user profiling strategies for search engine personalization, and observed the following problems in existing strategies.

Most personalization methods focused on the creation of one single profile for a user and applied the same profile to all of the user's queries. We believe that different queries from a user should be handled differently because a user's preferences may vary across queries. For example, a user who prefers information about fruit on the query "orange" may prefer the information about Apple Computer for the query "apple."

Existing click through-based user profiling strategies can be categorized into document-based and concept based approaches. They both assume that user clicks can be used to infer users' interests, although their inference methods and the outcomes of the inference are different. Document-based profiling methods try to estimate users' document preferences (i.e., users are interested in some documents more than others). On the other hand, concept based profiling methods aim to derive topics or concepts that users are highly interested. Document-based methods that consider both users' positive and negative preferences, to the best of our knowledge, there are no concept-based methods that considered both positive and negative preferences in deriving user's topical interests.

Most existing user profiling strategies only consider documents that users are interested in (i.e., users' positive preferences) but ignore documents that user's dislike (i.e., users' negative preferences). For example, if a user is interested in "apple" as a fruit, he/she may be interested specifically in apple recipes, but less interested in information about growing apples, while absolutely not interested in information about the company Apple Computer. In this case, a good user profile should favor information about apple recipes, slightly favor information about growing apple, while downgrade information about Apple Computer. Profiles built on both positive and negative user preferences can represent user interests at finer details.

The main contributions of this paper are:

- The query-oriented, concept-based user profiling method [11] to consider both users' positive and negative preferences in building users profiles. We proposed six user profiling methods that exploit a user's positive and negative preferences to produce a profile for the user using a Ranking SVM (RSVM).

- While document-based user profiling methods pioneered by Joachim's [10] capture users' document preferences (i.e., users consider some documents to be more relevant than others), our methods are based on users' concept preferences (i.e., users consider some topics/concepts to be more relevant than others).

- RSVM to learn from concept preferences weighted concept vectors representing concept-based user profiles. The weights of the vector elements, which could be positive or negative, represent the interestingness (or uninteresting ness) of the user on the concepts [11]. The weights that represent a user's interests are all positive, meaning that the method can only capture user's positive preferences.

II. AN OVERVIEW OF RELARED WORK

User profiling strategies can be broadly classified into two main approaches: document-based and concept-based approaches. Document-based user profiling methods aim at capturing users' clicking and browsing behaviors. Users' document preferences are first extracted from the click through data, and then, used to learn the user behavior model which is usually represented as a set of weighted features. On the other hand, concept-based user profiling methods aim at capturing users' conceptual needs. Users' browsed documents and search histories are automatically mapped into a set of topical categories. User profiles are created based on the users' preferences on the extracted topical categories.

A. Document-Based Methods

Most document-based methods focus on analyzing users' clicking and browsing behaviors recorded in the users' click through data. On Web search engines, click through data are important implicit feedback mechanism from users. An example of click through data for the query "apple," which contains a list of ranked search results presented to the user, with identification on the results that the user has clicked on. Several personalized systems that employ click through data to capture users' interest have been proposed [1], [2], [10].

B. Concept-Based Methods

Most concept-based methods automatically derive users' topical interests by exploring the contents of the users' browsed documents and search histories. Liu et al. [13] proposed a user profiling method based on users' search history and the Open Directory Project (ODP) [16]. The user profile is represented as a set of categories, and for each category, a set of keywords with weights. The categories stored in the user profiles serve as a context to disambiguate user queries. If a profile shows that a user is interested in certain categories, the search can be narrowed down by providing suggested results according to the user's preferred categories.

III. PERSONALIZED CONCEPT-BASED QUERY CLUSTERING

Our personalized concept-based clustering method consists of three steps. First, concept extraction algorithm, extract concepts and their relations from the Web-snippets returned by the search engine. Second, seven different concept-based user profiling strategies, to create concept based user profiles. Finally, the concept-based user profiles are compared with each other and against as baseline our previously proposed personalized concept-based clustering algorithm.

A. Concept Extraction

After a query is submitted to a search engine, a list of Web snippets is returned to the user. We assume that if a keyword/phrase exists frequently in the Web-snippets of a particular query, it represents an important concept related to the query because it coexists in close proximity with the query in the top documents. Thus, we employ the following support formula, which is inspired by the well-known problem of finding frequent item sets in data mining [7], to measure the interestingness of a particular keyword/phrase extracted from the Web-snippets.

B. Query Clustering Algorithm

Concept-based clustering algorithm with which ambiguous queries can be classified into different query clusters. Concept-based user profiles are employed in the clustering process to achieve personalization effect. First, a query-concept bipartite graph G is constructed by the clustering algorithm in which one set of nodes corresponds to the set of users' queries and the other corresponds to the sets of extracted concepts. Each individual query submitted by each user is treated as an individual node in the bipartite graph by labeling each query with a user identifier. Concepts with interestingness weights greater than zero in the user profile are linked to the query with the corresponding interestingness weight in G. Second, a two-step personalized clustering algorithm is applied to the bipartite graph G, to obtain clusters of similar queries and similar concepts.

$$sim(x, y) = \frac{N_x \cdot N_y}{\| N_x \| \| N_y \|}, \qquad (7)$$

where Nx is a weight vector for the set of neighbor nodes of node x in the bipartite graph G, the weight of a neighbor node nx in the weight vector Nx is the weight of the link connecting x and nx in G, Ny is a weight vector for the set

of neighbor nodes of node y in G, and the weight of a neighbor node ny in Ny is the weight of the link connecting y and ny in G.

Algorithm 1. Personalized Agglomerative Clustering

Input: A Query-Concept Bipartite Graph G

Output: A Personalized Clustered Query-Concept Bipartite

Graph Gp

// Initial Clustering

1:   Obtain the similarity scores in G for all possible pairs of query nodes using Equation (7).

2:   Merge the pair of most similar query nodes (qi,qj) that does not contain the same query from different users. Assume that a concept node c is connected to both query nodes qi and qj with weight wi and wj, a new link is created between c and (qi; qj) with weight w = wi +wj.

3:   Obtain the similarity scores in G for all possible pair's of concept nodes using Equation (7).

4:   Merge the pair of concept nodes (ci,cj) having highest similarity score. Assume that a query node q is connected to both concept nodes ci and cj with weight wi and wj, a new link is created between q and (ci; cj) with weight w = wi +wj.

5.   Unless termination is reached, repeat Steps 1-4.
     // Community Merging

6:   Obtain the similarity scores in G for all possible pairs of query nodes using Equation (7).

7:   Merge the pair of most similar query nodes (qi,qj) that contains the same query from different users. Assume that a concept node c is connected to both query nodes qi and qj with weight wi and wj, a new link is created between c and (qi; qj) with weight w = wi + wj.

8.   Unless termination is reached, repeat Steps 6-7.

IV. USER PROFILE STRATEGIES

Six user profiling strategies which are both concept-based and utilize users' positive and negative preferences. They are PJoachims_C, PmJoachims_C, PSpyNB_C, PClickþJoachims_C, PClickþmJoachims_C, and PClickþSpyNB_C.

A.Click-Based Method (PClick)

The concepts extracted for a query q using the concept extraction method the possible concept space arising from the query q. The concept

Space may cover more than what the user actually wants. For example, when the user searches for the query "apple," the concept space derived from our concept extraction method contains the concepts "macintosh," "ipod," and "fruit." If the user is indeed interested in "apple" as a fruit and clicks on pages containing the concept "fruit," the user Profile represented as a weighted concept vector should record the user interest on the concept "apple" and its neighborhood (i.e., concepts which having similar meaning as "fruit"), while downgrading unrelated concepts such as "macintosh," "ipod," and their neighborhood.

B.Joachims-C Method (PJoachims_C)

Given a list of search results for an input query q, if a user clicks on the document dj at rank j, all the concepts C(di) in the unclicked documents di above rank j are considered as less relevant than the concepts C(dj) in the document dj, i.e., (C(dj) <r' C(di)), where r' is the user's preference order of the concepts extracted from the search results of the query q).

C. mJoachims-*C Method (PmJoachims_C)*

Given a set of search results for a query, if document di at rank i is clicked, dj is the next clicked document right after di (no other clicked links between di and dj), and document dk at rank k between di and dj (i < k < j) is not clicked, then conceptsC(dk) in document dk is considered less relevant than the concepts C(dj) in document dj (C(dj) <r' C(dk)), where r' is the user's preference order of the concepts extracted from the search results of the query q).

D.SpyNB-C Method (PSpyNB_C)

Both Joachims and mJoachims are based on a rather strong assumption that pages scanned but not clicked by the user are considered uninteresting to the user, and hence, irrelevant to the user's query. the search engine context, most users would only click on a few documents (positive examples) that are relevant to them. Thus, only a limited number of positive examples can be used in the classification process, lowering the reliability of the predicted negative examples.

E.Click+Joachims-C Method (PClickþJoachims_C)

Integrate the click-based method, which captures only positive preferences, with the Joachims-C method, with which negative preferences can be obtained. We found that Joachims-C is good in predicting users' negative preferences. Since both the user profiles PClick and PJoachims_C are represented as weighted concept vectors, the two vectors can be combined.

F.Click+mJoachims-C Method (PClickþmJoachims_C)

Similar to Click+Joachims-C method, a hybrid method which combines PClick and PmJoachims_C is proposed.

G. *Click+SpyNB-C Method (PClickþSpyNB_C)*

Similar to Click+Joachims-C and Click+mJoachims-C methods.

V. EXPERIMENTAL RESULTS

We evaluate and analyze the seven conceptbased user profiling strategies (i.e., PClick, PJoachims_C, PmJoachims_C, PSpyNB_C, PClick+Joachims_C, PClick+mJoachims_C, and PClick+SpyNB_C). The seven concept-based user profiling strategies are compared using our personalized concept-based clustering algorithm [11]. The collected clickthrough data are used by the proposed user profiling strategies to create user profiles. The performance of a heuristic for determining the termination points of initial clustering and community merging based on the change of intracluster similarity. We show that user profiling methods that incorporate negative concept weights return termination points that are very close to the optimal points obtained by exhaustive search.

A. Experimental Setup

The query and click through data for evaluation are adopted from our previous work [11]. To evaluate the performance of our user profiling strategies, we developed a middleware for Google to collect click through data. We used 500 test queries, which are intentionally designed to have ambiguous. The clusters obtained from the algorithms are compared against the standard clusters to check for their correctness. The 100 users are invited to use our middleware to search for the answers of the 500 test queries (accessible at [3]). To avoid any bias, the test queries are randomly selected from 10 different categories.

The user profiles are employed by the personalized clustering method to group similar queries together according to users' needs. The personalized clustering algorithm is a two-phase algorithm which composes of the initial clustering phase to cluster queries within the scope of each user, and then, the community merging phase to group queries for the community.

B. Comparing Concept Preference Pairs Obtained Using Joachims-C, mJoachims-C, and SpyNB-C Methods

In this section, we evaluate the pair wise agreement between the concept preferences extracted using Joachims-C, mJoachims- C, and SpyNB-C methods. The three methods are employed to learn the concept preference pairs from the collected click through data. The learned concept preference pairs from different methods are manually evaluated by human evaluators to derive the fraction of correct preference pairs. We discard all the ties in the resulted concept preference pairs (i.e., pairs with the same concepts) to avoid ambiguity (i.e., both $c_i > c_j$ and $c_j > c_i$ exist) in the evaluation. RSVM is then employed to learn user profiles from the concept preference pairs.

C. Comparing PClick, PJoachims_C, PmJoachims_C, PSpyNB_C, PClick+Joachims_C, PClick+mJoachims_C, and PClick+SpyNB_C

An important observation from these three figures is that even though PJoachims_C, PmJoachims_C, and PSpyNB_C are able to capture users' negative preferences, they yield worse precision and recall ratings comparing to PClick. This is attributed to the fact that PJoachims_C, PmJoachims_C, and PSpyNB_C share a common deficiency in capturing users' positive preferences. A few wrong positive predictions would significantly lower the weight of a positive concept.

For example, assume that a positive concept c_i has been clicked many times, a preference $c_j <_{r0} c_i$ can still be generated by Joachims/mJoachims propositions, if there ever exists one case in which the user did not click on c_i but clicked on another document that was ranked lower in the result list. Since PJoachims_C, PmJoachims_C, and PSpyNB_C cannot effectively capture users' positive preferences, they perform worse than the baseline method PClick. On the other hand, PClick captures positive preferences based on user clicks, so an erroneous click made by users has little effect on the final outcome as long as the number of erroneous clicks is much less than that of correct clicks.

D. Termination Points for Individual Clustering to Community Merging

As initial clustering is run, a tree of clusters will be built along the clustering process. The termination point for initial clustering can be determined by finding the point at which the cluster quality has reached its highest (i.e., further clustering steps would decrease the quality). The same can be done for determining the termination point for community merging.

Fig. 1. Change in similarity values when performing personalized clustering using PClick+Joachims_C.
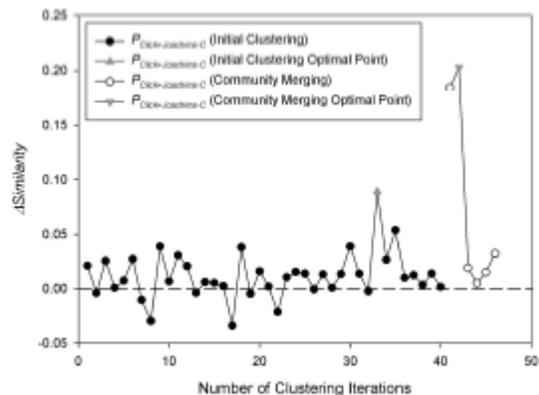
Fig. 2. Change in similarity values when performing personalized clustering using PClick+mJoachims_C.
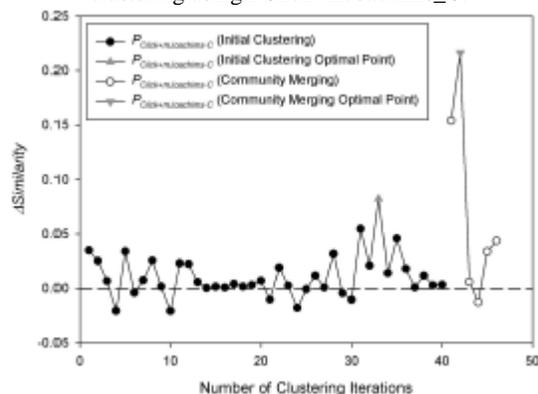
VI. CONCLUSION

An accurate user profile can greatly improve a search engine's performance by identifying the information needs for individual users. In this paper, we proposed and evaluated several user profiling strategies. The techniques make use of click through data to extract from Web-snippets to build concept-based user profiles automatically. We applied preference mining rules to infer not only users' positive preferences but also their negative preferences, and utilized both kinds of preferences in deriving user's profiles. The user profiling strategies were evaluated and compared with the personalized query clustering method that we proposed previously. Our experimental results show that profiles capturing both of the user's positive and negative preferences perform the best among the user profiling strategies studied. Apart from improving the quality of the resulting clusters, the negative preferences in the proposed user profiles also help to separate similar and dissimilar queries into distant clusters, which help to determine near optimal terminating points for our clustering algorithm.

We plan to take on the following two directions for future work. First, relationships between users can be mined from the concept-based user profiles to perform collaborative filtering. This allows users with the same interests to share their profiles. Second, the existing user profiles can be used to predict the intent of unseen queries, such that when a user submits a new query, personalization can benefit the unseen query. Finally, the concept-based user profiles can be integrated into the ranking algorithms of a search engine so that search results can be ranked according to individual users' interests.

REFERENCES

[1]   E. Agichtein, E. Brill, and S. Dumais, "Improving Web Search Ranking by Incorporating User Behavior Information," Proc. ACM SIGIR, 2006.
[2]   E. Agichtein, E. Brill, S. Dumais, and R. Ragno, "Learning User Interaction Models for Predicting Web Search Result Preferences," Proc. ACM SIGIR, 2006.
[3]   Appendix: 500 Test Queries, http://www.cse.ust.hk/~dlee/ tkde09/Appendix.pdf, 2009.
[4]   R. Baeza-yates, C. Hurtado, and M. Mendoza, "Query Recommendation Using Query Logs in Search Engines," Proc. Int'l Workshop Current Trends in Database Technology, pp. 588-596, 2004.
[5]   D. Beeferman and A. Berger, "Agglomerative Clustering of a Search Engine Query Log," Proc. ACM SIGKDD, 2000.
[6]   C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender, "Learning to Rank Using Gradient Descent," Proc. Int'l Conf. Machine learning (ICML), 2005.
[7]   K.W. Church, W. Gale, P. Hanks, and D. Hindle, "Using Statistics in Lexical Analysis," Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon, Lawrence Erlbaum, 1991.
[8]   Z. Dou, R. Song, and J.-R. Wen, "A Largescale Evaluation and Analysis of Personalized Search Strategies," Proc. World Wide Web (WWW) Conf., 2007.
[9]   S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing," ACM Web Intelligence and Agent System, vol. 1, nos. 3/4, pp. 219-234, 2003.
[10]   T. Joachims, "Optimizing Search Engines Using Clickthrough Data," Proc. ACM SIGKDD, 2002.
[11]   K.W.-T. Leung, W. Ng, and D.L. Lee, "Personalized Concept-Based Clustering of Search Engine Queries," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 11, pp. 1505-1518, Nov. 2008.
[12]   B. Liu, W.S. Lee, P.S. Yu, and X. Li, "Partially Supervised Classification of Text Documents," Proc. Int'l Conf. Machine Learning (ICML), 2002.
[13]   F. Liu, C. Yu, and W. Meng, "Personalized Web Search by Mapping User Queries to Categories," Proc. Int'l Conf. Information and Knowledge Management (CIKM), 2002.
[14]   Magellan, http://magellan.mckinley.com/, 2008.
[15]   W. Ng, L. Deng, and D.L. Lee, "Mining User Preference Using Spy Voting for Search Engine Personalization," ACM Trans. Internet Technology, vol. 7, no. 4, article 19, 2007.
[16]   Open Directory Project, http://www.dmoz.org/, 2009.
[17]   M. Speretta and S. Gauch, "Personalized Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence, 2005.
[18]   Q. Tan, X. Chai, W. Ng, and D. Lee, "Applying Co-training to Clickthrough Data for Search Engine Adaptation," Proc. Database Systems for Advanced Applications (DASFAA) Conf., 2004.
[19]   J.-R. Wen, J.-Y. Nie, and H.-J. Zhang, "Query Clustering Using User Logs," ACM Trans. Information Systems, vol. 20, no. 1, pp. 59- 81, 2002.
[20]   Y. Xu, K. Wang, B. Zhang, and Z. Chen, "Privacy-Enhancing Personalized Web Search," Proc. World Wide Web (WWW) Conf., 2007.

Authors

Mr.Kokila.R received B.E degree in CSE from Anna University, Chennai and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. Her research interest Includes Computer Networks and Data Mining.

Mrs.N.Krishnammal received the B.E degree in ECE from Anna University, Chennai and Received the M.E degree in CSE from Anna University of technology, Coimbatore and pursuing Phd in Networks under Anna university of technology, Coimbatore. She is currently working as Assistant Professor in Department of CSE in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Her main research interest is Computer Networks.

COMPARISION OF VARIOUS NOISE REMOVALS USING BAYESIAN FRAMEWORK

Er. Ravi Garg[1], Er. Abhijeet Kumar[2]

(Student, ECE Department, M. M. Engg. College, Mullana)[1]
(Lecturer, ECE Department, M. M. Engg. College, Mullana)[2]

ABSTRACT:
A noise is introduced in the transmission medium due to a noisy channel, errors during the measurement process and during quantization of the data. For digital storage each element in the imaging chain such as lenses, film, digitizer, etc. contributes to the degradation. Image noise removal is often used in the field of photography or publishing where an image was somehow degraded but needs to be improved before it can be printed. This paper reviews the Bayesian Estimation process for statistical signal processing. Different noise models including additive and multiplicative types are used. They include Gaussian noise, salt and pepper noise, speckle noise and Poisson noise. Selection of the denoising algorithm is application dependent. Hence, it is necessary to have knowledge about the noise present in the image so as to select the appropriate noise removal algorithm. The filtering approach has been proved to be the best when the image is corrupted with salt and pepper noise. The wavelet based approach finds applications in denoising images corrupted with Gaussian noise. In the case where the noise characteristics are complex, the multifractal approach can be used. Bayesian estimation process is used to optimize the removal of Poisson noise. A quantitative measure of comparison is provided by the signal to noise ratio of the image.

KEYWORDS: Bayesian estimator, prior Distribution, Posterior Distribution, Likelihood, Gaussian, salt and pepper, speckle, Poisson Noise.

1. INTRODUCTION

Visual information transmitted in the form of digital images is becoming a major method of communication in the modern age, but the image obtained after transmission is often corrupted with noise. The received image needs processing before it can be used in applications. Image noise removal involves the manipulation of the image data to produce a visually high quality image. For this type of application we need to know something about the degradation process in order to develop a model for it. When we have a model for the degradation process, the inverse process can be applied to the image to restore it back to the original form. This type of image restoration is often used in space exploration to help eliminate artifacts generated by mechanical jitter in a spacecraft or to compensate for distortion in the optical system of a telescope. Image denoising finds applications in fields such as astronomy where the resolution limitations are severe, in medical imaging where the physical requirements for high quality imaging are needed for analyzing images of unique events, and in forensic science where potentially useful photographic evidence is sometimes of extremely bad quality.

Nonlinear filtering is the process of estimating and tracking the state of a nonlinear stochastic system from non-Gaussian noisy observation data. In this technical memorandum, we present an overview of techniques for nonlinear filtering for a wide variety of conditions on the nonlinearities and on the noise. We begin with the development of a general Bayesian approach to filtering which is applicable to all linear or nonlinear stochastic systems. We show how Bayesian filtering requires integration over probability density functions that cannot be accomplished in closed form for the general nonlinear, non-Gaussian multivariate system.

2. ADDITIVE AND MULTIPLICATIVE NOISES

Noise is undesired information that contaminates the image. In the image denoising process, information about the type of noise present in the original image plays a significant role. Typical images are corrupted with noise modeled with either a Gaussian, uniform, or salt or pepper distribution. Another typical noise is a speckle noise, which is multiplicative in nature. The behavior of each of these noises is described below. The digital image acquisition process converts an optical image into a continuous electrical signal that is, then, sampled. At every step in the process there are fluctuations caused by natural phenomena, adding a random value to the exact brightness value for a given pixel.

2.1 GAUSSIAN NOISE

Gaussian noise is evenly distributed over the signal. This means that each pixel in the noisy image is the sum of the true pixel value and a random Gaussian distributed noise value. As the name indicates, this type of noise has a Gaussian distribution, which has a bell shaped probability distribution function given by,

$$F(g) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(g-m)^2/2\sigma^2},$$

(1)

Where *g* represents the gray level, *m* is the mean or average of the function and σ is the standard deviation of the noise. Graphically, it is represented as shown in Fig 1. The original image is shown in Fig 2 and the image after Gaussian Noise (variance = 0.05) addition is shown in Fig 3.
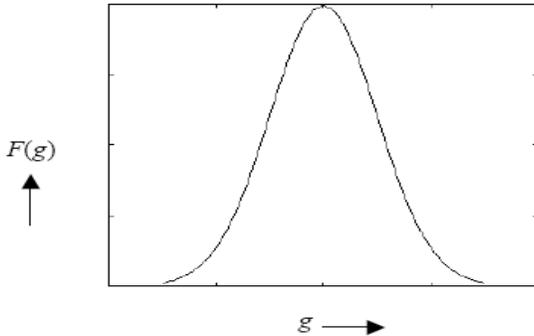

Fig1: PDF for Gaussian Noise

Gaussian noise can be reduced using a spatial filter. However, it must be kept in mind that when smoothing an image, we reduce not only the noise, but also the fine-scaled image details because they also correspond to blocked high frequencies. The most effective basic spatial filtering techniques for noise removal include: mean filtering, median filtering and Gaussian smoothing Crimmins Speckle Removal filter can also produce good noise removal. More sophisticated algorithms which utilize statistical properties of the image and/or noise fields exist for noise removal. For example, adaptive smoothing algorithms may be defined which adjust the filter response according to local variations in the statistical properties of the data.
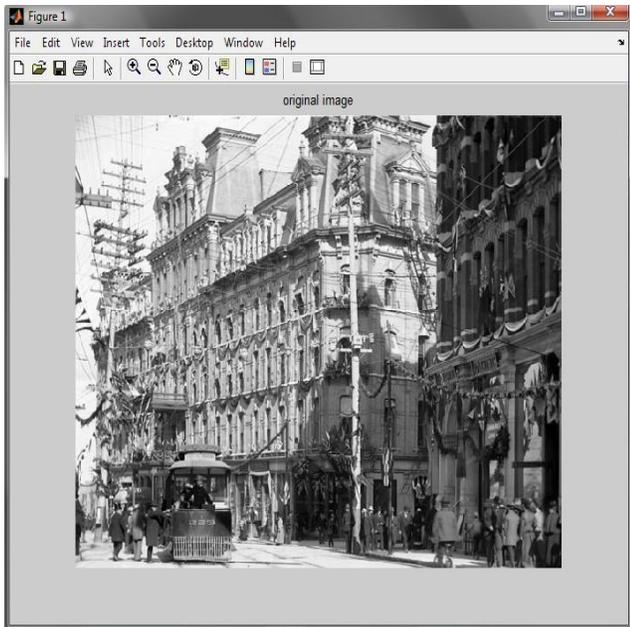
Fig 2: Original Image

Fig 3: Image after Gaussian Noise addition
(Noise Variance = 0.05)

2.2 SALT AND PEPPER NOISE

For this kind of noise, conventional low pass filtering, e.g. mean filtering or Gaussian smoothing is relatively unsuccessful because the corrupted pixel value can vary significantly from the original and therefore the mean can be significantly different from the true value. A median filter removes drop-out noise more efficiently and at the same time preserves the edges and small details in the image better. Conservative smoothing can be used to obtain a result which preserves a great deal of high frequency detail, but is only effective at reducing low levels of noise. In salt and pepper noise (sparse light and dark disturbances), pixels in the image are very different in color or intensity from their surrounding pixels; the defining characteristic is that the value of a noisy pixel bears no relation to the color of surrounding pixels. Generally this type of noise will only affect a small number of image pixels. When viewed, the image contains dark and white dots, hence the term salt and pepper noise. Typical sources include flecks of dust inside the camera and overheated or faulty CCD elements.

Salt and pepper noise is an impulse type of noise, which is also referred to as intensity spikes. This is caused generally due to errors in data transmission. It has only two possible values, *a* and *b.* The probability of each is typically less than 0.1. The corrupted pixels are set alternatively to the minimum or to the maximum value, giving the image a "salt and pepper" like appearance. Unaffected pixels remain unchanged. For an 8-bit image, the typical value for pepper noise is 0 and for salt noise 255. The salt and pepper noise is generally caused by malfunctioning of pixel elements in the camera sensors, faulty memory locations, or timing errors in the digitization process. The probability density function for this type of noise is shown in Fig 4. Salt and pepper noise with a variance of 0.05 is shown in Fig 5.
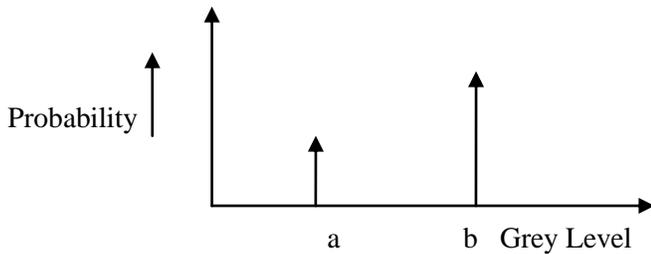
Fig 4: PDF for Salt and Pepper Noise


Fig 6: Gamma Distribution

Speckle noise is a granular noise that inherently exists in and degrades the quality of the active radar and synthetic aperture radar (SAR) images. Speckle noise in conventional radar results from random fluctuations in the return signal from an object that is no bigger than a single image-processing element. It increases the mean grey level of a local area. There are many forms of adaptive speckle filtering, including the Lee filter, the Frost filter, and the Refined Gamma Maximum-A-Posteriori (RGMAP) filter. They all rely upon three fundamental assumptions in their mathematical models, however:

1. Speckle noise in SAR is a multiplicative noise, i.e. it is in direct proportion to the local grey level in any area.

2. The signal and the noise are statistically independent of each other.

3. The sample mean and variance of a single pixel are equal to the mean and variance of the local area.

Fig 5: Image after Salt and Pepper Noise addition (Noise Variance = 0.05)

2.3 SPECKLE NOISE

Speckle noise is a multiplicative noise. This type of noise occurs in almost all coherent imaging systems such as laser, acoustics and SAR (Synthetic Aperture Radar) imagery. The source of this noise is attributed to random interference between the coherent returns. Fully developed speckle noise has the characteristic of multiplicative noise. Speckle noise follows a gamma distribution and is given as,

$$F(g) = \frac{g^{\alpha-1}}{(\alpha-1)! \, a^{\alpha}} e^{-\frac{g}{a}},$$

(2)

where variance is $a^2\alpha$ and g is the gray level. The gamma distribution is given below in Fig 6. On an image, speckle noise (with variance 0.05) looks as shown in Fig 7.

Fig 7: Image after Speckle Noise addition (Noise Variance = 0.05)

2.4 POISSON NOISE

The dominant noise in the lighter parts of an image from an image sensor is typically that caused by statistical quantum fluctuations, that is, variation in the number of photons sensed at a given exposure level; this noise is known as photon shot noise or Poisson Noise. Poisson noise has a root-mean-square value proportional to the square root of the image intensity, and the noises at different pixels are independent of one another. Poisson noise follows a Poisson distribution, which is usually not very different from Gaussian. Poisson noise is a type of electronic noise that may be dominant when the finite number of particles that carry energy (such as electrons in an electronic circuit or photons in an optical device) is sufficiently small so that uncertainties due to the Poisson distribution, which describes the occurrence of independent random events, are of significance. It is important in electronics, telecommunications, optical detection, and fundamental physics. The magnitude of shot noise increases according to the square root of the expected number of events, such as the electrical current or intensity of light. But since the strength of the signal itself increases more rapidly, the relative proportion of Poisson noise decreases and the signal to noise ratio (considering only Poisson noise) increases anyway. Thus Poisson noise is more frequently observed with small currents or light intensities following sufficient amplification. Since the standard deviation of Poisson noise is equal to the square root of the average number of events N, the signal-to-noise ratio is given by:

$$\mathrm{SNR} = N/\sqrt{N} = \sqrt{N} \qquad (3)$$

Thus when N is very large, the signal-to-noise ratio is very large as well.
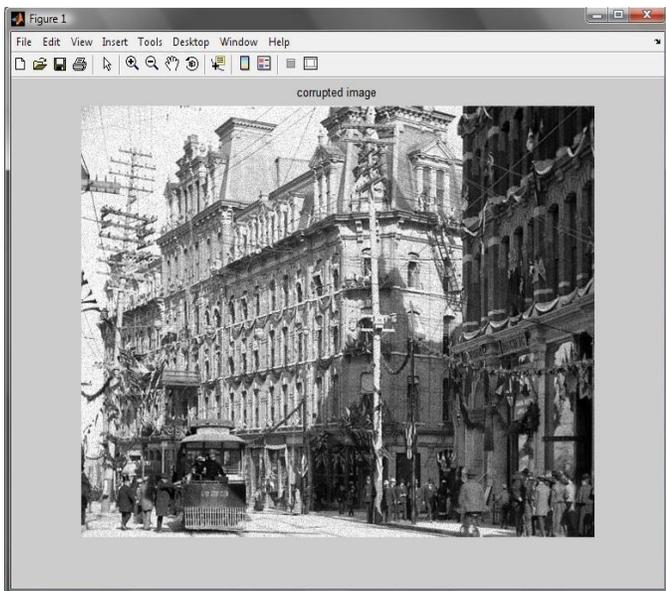
Fig 8: Image after Poisson Noise addition
(Noise Variance = 0.05)

3. BAYESIAN ESTIMATOR

Bayesian estimation is a framework for the formulation of statistical inference problems. In the prediction or estimation of a random process from a related observation signal, the Bayesian philosophy is based on combining the evidence contained in the signal with prior knowledge of the probability distribution of the process. Bayesian methodology includes the classical estimators such as maximum a posteriori (MAP), maximum-likelihood (ML), minimum mean square error (MMSE) and minimum mean absolute value of error (MAVE) as special cases. The hidden Markov model, widely used in statistical signal processing, is an example of a Bayesian model. Bayesian inference is based on minimization of the so-called Baye's risk function, which includes a posterior model of the unknown parameters given the observation and a cost-of-error function.

Estimation theory is concerned with the determination of the best estimate of an unknown parameter vector from an observation signal, or the recovery of a clean signal degraded by noise and distortion. For example, given a noisy sine wave, we may be interested in estimating its basic parameters (i.e. amplitude, frequency and phase), or we may wish to recover the signal itself. An estimator takes as the input a set of noisy or incomplete observations, and, using a dynamic model (e.g. a linear predictive model) and/or a probabilistic model (e.g. Gaussian model) of the process, estimates the unknown parameters. The estimation accuracy depends on the available information and on the efficiency of the estimator.

Bayesian theory is a general inference framework. In the estimation or prediction of the state of a process, the Bayesian method employs both the evidence contained in the observation signal and the accumulated prior probability of the process. Consider the estimation of the value of a random parameter vector θ, given a related observation vector y. From Baye's rule the posterior probability density function (pdf) of the parameter vector θ given y, $f(\theta/y)$ can be expressed as

$$f(\theta/y) = \frac{f(y/\theta)f(\theta)}{f(y)} \qquad (4)$$

Where for a given observation, f(y) is a constant and has only a normalizing effect. Thus there are two variable terms in the Equation 4. One term $f(y/\theta)$ is the likelihood that the observation signal y was generated by the parameter vector θ and the second term is the prior probability of the parameter vector having a value of θ. Conceptually Bayesian Estimator combines

1. The likelihood, i.e., the data, with

2. The prior

3.1 DYNAMIC AND PROBABILITY MODELS IN ESTIMATION

Optimal estimation algorithms utilize dynamic and statistical models of the observation signals. A dynamic predictive model captures the correlation structure of a signal, and models the dependence of the present and future values of the signal on its past trajectory and the input stimulus. A statistical probability model characterizes the random fluctuations of a signal in terms of its statistics, such as the mean and the covariance, and most completely in terms of a probability model. As an illustration consider the estimation of a *P*-dimensional parameter vector $\theta = [\theta_0, \theta_1, ... \theta_{p-1}]$ from a noisy observation vector $y=[y(0), y(1), ..., y(N–1)]$ modeled as

$$y = h(\theta, \mathbf{x}, \mathbf{e}) + \boldsymbol{n} \qquad (5)$$

In Fig 9, the distributions of the random noise $\boldsymbol{n}$, the random input e and the parameter vector θ are modeled by probability density functions, $f(\boldsymbol{n})$, $f(\boldsymbol{e})$, and $f(\theta)$ respectively. The pdf model most often used is the Gaussian model. Predictive and statistical models of a process guide the estimator towards the set of values of the unknown parameters that are most consistent with both the prior distribution of the model parameters and the noisy observation. In general, the more modeling information used in an estimation process, the better the results, provided that the models are an accurate characterization of the observation and the parameter process.
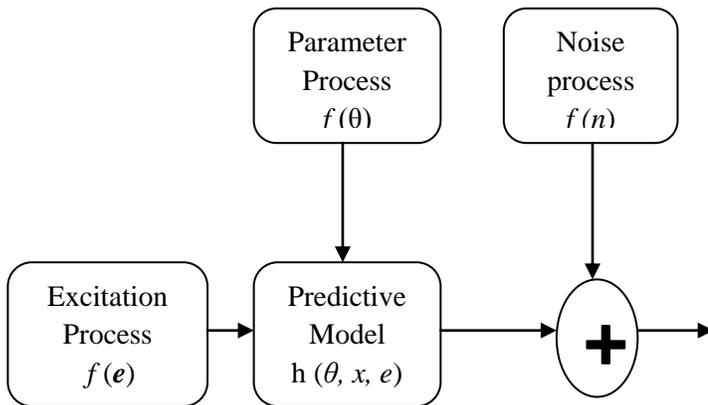
Fig 9. A random process **y** is described in terms of a predictive model h (·), and statistical models $f(e)$, $f(\theta)$ and $f(\boldsymbol{n})$

3.2 PARAMETER ESTIMATION AND SIGNAL RESTORATION

Parameter estimation and signal restoration are closely related problems. The main difference is due to the rapid fluctuations of most signals in comparison with the relatively slow variations of most parameters. For example, speech sounds fluctuate at speeds of up to 20 kHz, whereas the underlying vocal tract and pitch parameters vary at a relatively lower rate of less than 100

Hz. This observation implies that normally more averaging can be done in parameter estimation than in signal restoration. As a further example, consider the interpolation of a sequence of lost samples of signal given N recorded samples, as shown in Fig 10.

$$y = \boldsymbol{X}\,\theta + \mathbf{e} + \mathbf{n} \qquad (6)$$

Where y is the observation signal, $\boldsymbol{X}$ is the signal matrix, θ is the AR parameter vector, e is the random input of the AR model and $\boldsymbol{n}$ is the random noise. Using Equation 6, the signal restoration process involves the estimation of both the model parameter vector θ and the random input e for the lost samples. Assuming the parameter vector θ is time-invariant, the estimate of θ can be averaged over the entire N observation samples, and as N becomes infinitely large, a consistent estimate should approach the true parameter value.
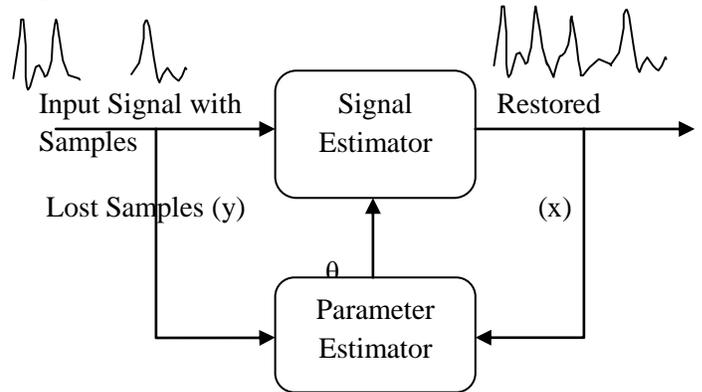

Fig 10 : Signal restoration using a parametric model of the signal process.

The difficulty in signal interpolation is that the underlying excitation e of the signal x is purely random and, unlike θ, it cannot be estimated through an averaging operation.

4. SNR:

Signal to Noise Ratio is probably the most well-known measure of them all. It is defined as the quotient between the signal and noise energy. If a signal is a scalar function $f(x)$, the energy is usually defined as:

$$E(f(x)) = \int (f(x))^2\,dx \qquad (7)$$

Consider an observed signal $X = S + N$, where S is the interesting part of the signal and N is the noise. SNR for X is usually defined as:

$$SNR(X) = 20\log_{10}\frac{E(S)}{E(N)} \text{ dB} \qquad (8)$$

The list of problems with SNR can be made long. The most obvious problem is, however, that it can usually not be directly measured. If S and N are known, measuring SNR is no problem, but then noise reduction will not be

needed. The noise free signal is already known. Various schemes of estimating SNR can be considered, but in doing so, we always need to decide what parts of the observed signal are interesting and what parts are noises.

5. RMSE

Another error measure of great importance is the Root Mean Square Error. RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{\sum i\ (xi - si)2}{N}} \qquad (9)$$

where xi is the sampled signal value at position i and s_i is the noise free value at the same position. It is obvious that if $x = s + n$, that is, if we add some uncorrelated noise n with zero mean to s and observe the resulting sum, RMSE will be equal to the standard deviation of the noise. RMSE is, in other words, an absolute measure of the noise amplitude. As with SNR, RMSE requires knowledge of the true noise free signal, which limits its use significantly.

6. RESULTS:

Comparison of SNR (Signal to Noise ratio) using Baye's Estimator (σ= 0.5):

	Gaussian Noise	Salt and Pepper Noise	Speckle Noise	Poisson Noise
SNR	22.2760	22.6158	24.8829	42.3882

Comparison of MSE (Mean Square Error) using Baye's Estimator (σ= 0.5):

	Gaussian Noise	Salt and Pepper Noise	Speckle Noise	Poisson Noise
MSE	0.2183	0.2061	0.1733	0.0327
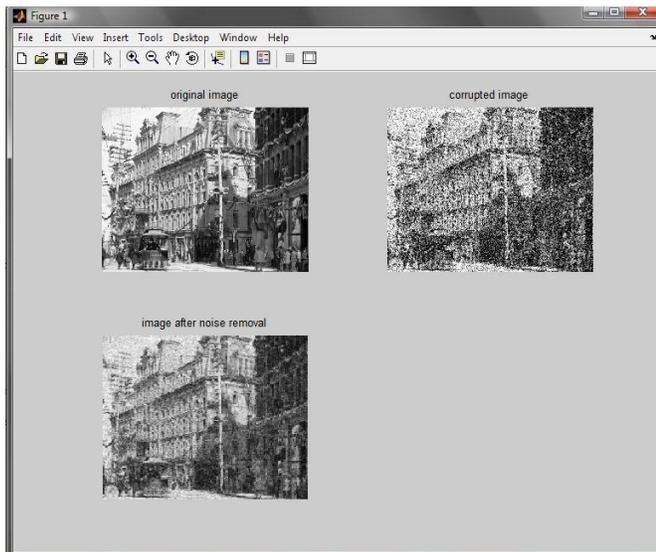
Fig 11: Removal of Poisson Noise by Bayesian estimator

7. CONCLUSION

The paper emphasizes on the SNR and MSE for various noises using Bayesian Estimator. The Results shows that Bayesian Estimator Optimizes the Poisson Noise removal as its Signal to Noise ratio (SNR) is maximum and least Mean square error (MSE). The SNR for Gaussian noise is minimum and also the MSE is maximum. Even if level of decomposition is increased the results becomes better.

REFERENCES:

[1] Z.Zhang, Y.Zhang, Y.Pan, L.Xiao, "A Bayes Thresholding Method Based on Edge Protection Strategies," IEEE Transactions On Multimedia technology, pp-12-15, August 2011.

[2] M.Zhou, H.Chen, J.Paisley, L.Ren, L.Li," Nonparametric Bayesian Dictionary Learning for Analysis of Noisy and Incomplete Images," IEEE Transaction on Image processing, pp-1-16, June 2011.

[3] S.M.Hashemi, S.Beheshti." Adaptive Image Denoising By Rigorous Bayesshrink Thresholding," IEEE Transaction on statistical Signal processing, pp-713-716, June, 2011.

[4] Joyee Ghosh and Merlise A. Clyde, "Rao-Blackwellization for Bayesian Variable Selection and Model Averaging in Linear and Binary Regression:A Novel Data Augmentation Approach" pp-1-36,2011.

[5] Zhixin Chen, "Simulation of Spectral Subtraction Based Noise Reduction Method" International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 2, No.8, pp-30-32, 2011.

[6] Tirza Routtenberg and Joseph Tabrikian, "Optimal Bayesian Parameter Estimation With Periodic Criteria" IEEE transaction on Sensor Array and Multichannel Signal Processing Workshop, pp-53-56, 2010.

[7] J.Harikiran, B.Saichandana, B.Divakar, "Impulse Noise Removal in Digital Images" International Journal of Computer Applications, Vol.10, pp-39-42, Nov 2010.

[8] Alexander Wong, Akshaya Mishra, Kostadinka Bizheva, and David A. Clausi, "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery" Vol. 18, No. 8, pp-8338-8352, 12 April 2010.

[9] Aliaa, A.A.Youssif, A.A.Darwish and A.M.M.Madbouly, "Adaptive Algorithm for Image Denoising Based on Curvelet Threshold" IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.1,pp-322-328 January 2010.

[10] Miika Toivanen and Jouko Lampinen, "Incremental Object Matching with Bayesian Methods and Particle Filters", IEEE Transaction on Digital Image Computing Techniques and Applications,pp-111-118, 2009.

Reheating Refrigeration System

Meemo Prasad
Assistant Engineer, SEPCO1

Abstract

The title "Reheating Refrigeration System" has the objective to utilize the rejected heat from the condenser of an air conditioner in an economy way.

This will be done by adding an arrangement called "REHEATER SYSTEM" which will have "INSULATED CONDUCTOR" whose one end is in the "HEAT EXCHANGER" at place before the condenser and the other end is in the "REHEATER" at a place after the evaporator and just before the compressor of the refrigeration system. What amount of heat going to the condenser, will be transferred to the reheater by the conductor as there will be some temperature difference in the heat exchanger and the reheater. Owing to this, the temperature and pressure will increase before the compressor in the reheater. Thus the compressor work can be reduced to minimum and the refrigerating effect will tend to increase. So mainly three benefits can be achieved as given as:

1. Utilization of wasting heat.

2. Reduction in work input to the compressor.

3. Better refrigerating effect

Compressor	Discharge Temperature
Screw Compressor (Indirect cooled)	70 → 80 °C
Screw Compressor (Injection cooled)	50 → 60°C
Reciprocating Compressors	85 → 110°C
Boosters (Rotaries & Reciprocating)	75 → 85°C

1. WHY IT IS NEEDED

In the vapour compression refrigeration system (VCRS), for an air conditioner generally a large amount of heat is generated in the refrigerant after its compression in the compressor and that is rejected to the atmosphere through the condenser and thus the heat is wasted. With this, the compressor requires higher power to operate when the cooling rate is needed more. Thus for best practices we need to adapt a project.

It is seen that the discharge gas coming out from the compressor is in a superheated state and some heat can be recovered from this refrigerant gas by de-superheating it before it enters to the condenser. The discharge temperatures in most refrigeration systems are quite high (in the range of 70°C to 100°C).

The superheat can be recycled to heat the refrigerant coming from the evaporator, before the compressor and can be used to increase its pressure thus can reduce the compressor work input to achieve the desired cooling effect.
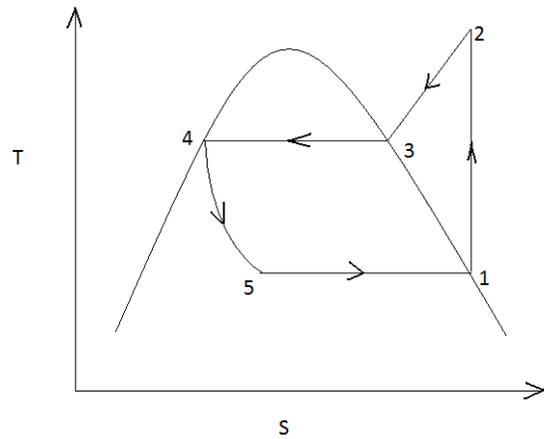
Fig.1 T-s diagram of the VCRS

The discharge temperatures of various types of compressors for refrigerant F22 are generally as follows:

2. REHEATING REFERIGERATION SYSTEM (RRS)

To overcome with this problem, the "Reheating System" is added to the refrigeration system, thus naming Reheating Refrigeration System.

In the reheating system, mainly three components are there;

a) Heat exchanger,

b) Reheater, and

c) Insulated conductor.

The reheater is placed after the evaporator and before the compressor and the heat exchanger is placed after the compressor and before the condenser. Insulated conductor is connecting to both reheater and heat exchanger.
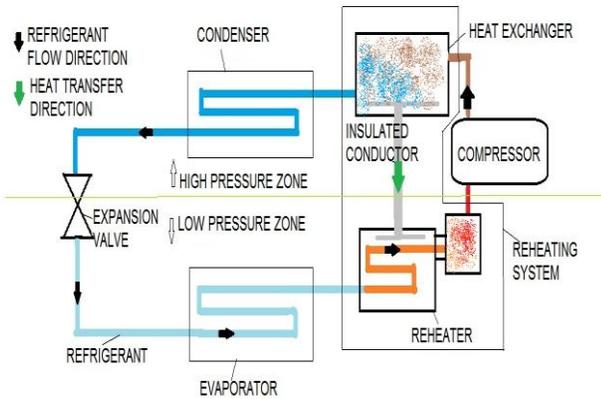

Fig.2 Reheating Refrigeration System
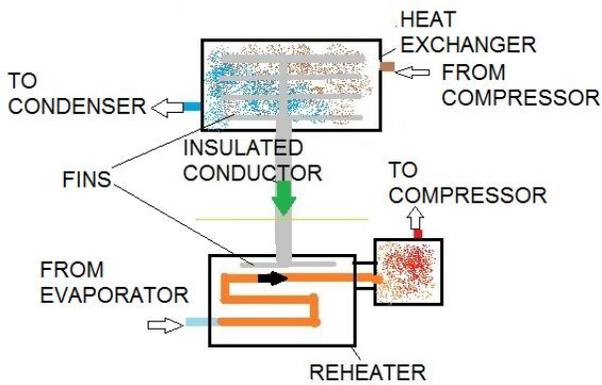
Fig. 3 Reheating system

The heat exchanger is that device which facilitates the heat transfer from the compressed refrigerant coming in it, with the fins of the conductor, which is within it, by the conduction mode. The heat exchanger adapted is a box type device in which the refrigerant only flows and transfers the heat due to temperature difference.

The heat exchanger focuses to increase heat exchange rate. If the fins are achieving higher heat from the refrigerant, means, will be in higher temperature and the refrigerant looses maximum heat to the fins. Then the effectiveness of heat transfer will be Maximum. For better storing of the refrigerant in the heat exchanger, for higher temperature achievement, it should contain suitable material at inside, of higher thermal conductivity.
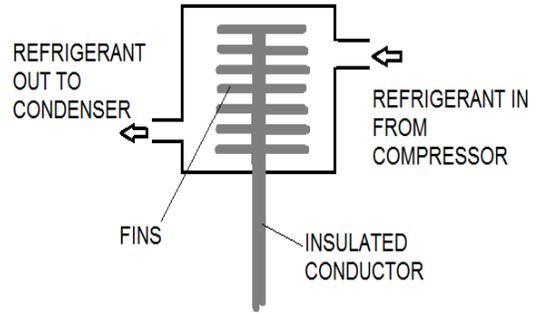
Fig. 4 Heat exchanger

The reheater is that device in which the heat obtained from the heat exchanger is utilized to raise the pressure of refrigerant coming in it, by incrementing the temperature. The refrigerant coming from the evaporator enters the reheater. In the reheater, the fins of the conductor, getting heat from refrigerant in the heat exchanger, transfers the heat to the refrigerant. Thus the temperature of refrigerant increases and the reheater is such designed that it will provide a constant volume system so the pressure would tend to increase. Now the pre-pressurised refrigerant will be forwarded to the compressor.
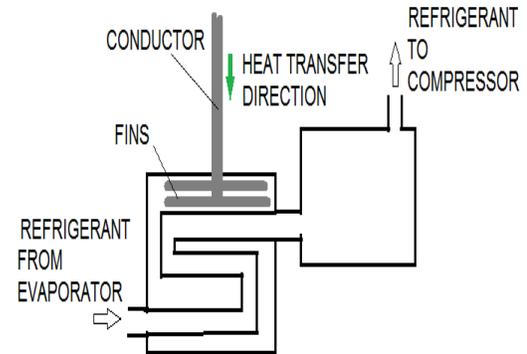
Fig.5 Reheater

The insulated conductor being used in the reheating refrigeration system, transfers the heat from the Heat exchanger to the Reheater. The conductor is insulated so that heat loss to the surroundings may be reduced. The conductor will have fins at its ends. Fins at one end will be inside the heat exchanger and the fins at other end will be inside the reheater. The material should of higher thermal conductivity like copper, aluminium etc.

4. ASSUMPTIONS MADE

i.   One dimensional heat flow

ii.   Steady state heat dissipation

iii. No internal heat generation

iv.   Homogeneous and isotropic material, thermal conductivity of the material is constant

v.   Uniform heat transfer coefficient on the entire surface

vi.   Thickness of the fin is small as compared to width and length

vii. Negligible radiation exchange with the surrounding.

5. THE PROCESS

When the heat is gained by the refrigerant in the evaporator, the refrigerant gets evaporated as it will have lower boiling point generally below 0°C and its density becomes lesser and it goes towards the compressor but as there is reheater before the compressor the refrigerant will come to the reheater .For the first stage, in the reheater the refrigerant will normally enter and goes out to the compressor. In the compressor the pressure of the refrigerant is increased with the temperature. Now the refrigerant enters the heat exchanger, where the heat of the compressed refrigerant is given to the conductor as it is having lower temperature (conductor is in contact with the temperature of refrigerant at the reheater). The rest amount of heat of refrigerant is rejected to the atmosphere in the condenser and it comes in the liquid phase. And it goes to the evaporator passing through the capillary tube where its pressure is reduced.

In the second stage, the refrigerant evaporated in the evaporator enters the reheater and affected by the heat transfer which is by the conductor getting heat from the heat exchanger and the conductor transfers the heat to the refrigerant. The vaporized refrigerant gets heated and its temperature and pressure is increased due the reheater is providing the constant volume system. Now the pressure of refrigerant is higher than that of the previous. This pre-pressurized vapour refrigerant reaches to the compressor where the compressor requires lesser power to get the predefined previous temperature. The refrigerant after getting predefined temperature goes to the heat exchanger where it loses the heat to the conductor again as previously and reaches to the condenser and the further process is similar as previous stage. Thus the cycle is completed.

The processes are:

6-1: heat addition at the constant temperature and pressure in the evaporator

1-2: heat addition at constant volume in the reheater

2-3: isentropic compression in the compressor

3-4: drop in temperature in the heat exchanger at constant pressure

4-5: heat rejection in the condenser at constant temperature and pressure

5-6: pressure drop in the throttle valve isenthalpically
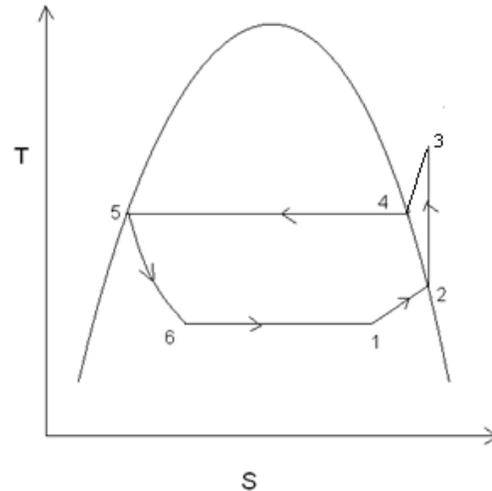
Fig.6 T-s diagram of reheating refrigeration system
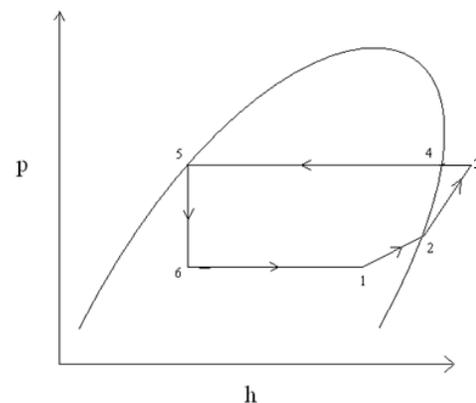
Fig.7 Reheating refrigeration system on p-h diagram

6. GENERAL REFRIGERATION SYSTEM VERSUS REHEATING REFRIGERATION SYSTEM

The cycle analysis of reheating refrigeration system with T-s diagram and p-h diagram are shown in the figures below

In the T-s diagram,

- For simple vapour compression refrigeration system (1):-

  Area under the process 1-2-3-4-5-1 is showing the work input to the system and the area 1-5-A-D-1 is showing the cooling effect of the system.

- For reheating refrigeration system (2):-

  Area under process 6-7-8-9-10-11-6 is giving the work input to the system and the area under 7-6-11-B-C-7 is showing the cooling effect of the system.

Here,

(Work input) $_1$ > (work input) $_2$

(Cooling effect) $_1$ < (cooling effect) $_2$
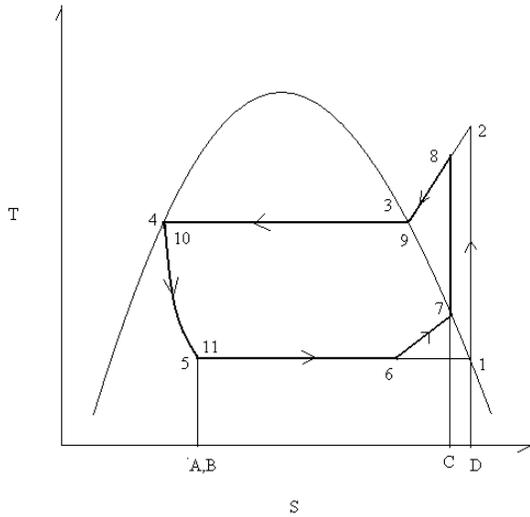
Fig.8 T-s diagram of reheating refrigeration system

In the p-h diagram,

- For simple vapour compression refrigeration system (1):-

Work input to the system = $h_2 - h_1$

And refrigerating effect = $h_1 - h_5$

- For reheating refrigeration system (2):-

- Work input to the system = $h_8 - h_7$

- And refrigerating effect = $h_7 - h_5$
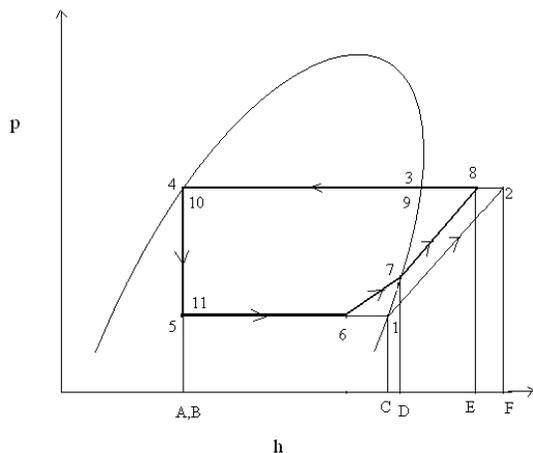
Fig. 9 p-h diagram of reheating refrigeration system

It is clear that

(Work input) $_1$ > (work input) $_2$

(Refrigerating effect) $_1$ < (refrigerating effect) $_2$

Thus it may be concluded that

Coefficient of performance of RRS > Coefficient of performance of simple VCRS

$$(\text{since COP} = \frac{\text{refrigerating effect}}{\text{work input}})$$

Therefore the reheating refrigeration system increases the coefficient of performance of the system by simply utilization of the heat wasting in the condenser and providing the economical benefits as low capacity compressor can also be used.

7. BENEFITS

With the adaptation of reheating refrigeration system the following advantages can be achieved:

1. There is saving of wasted heat in the condenser

2. The power consumption in the compressor is reduced to minimum

3. The amount of refrigerant may also be reduced

4. Better cooling effect can be achieved

5. The apparatus are easier to construction and easily available

6. There is no change in the components of the refrigeration system

7. There is no change in the air conditioning parameters.

8. LIMITATIONS

1. The life of the refrigerant may decrease faster as it will gain more heat

2. The distance between the heat exchanger and reheater should be adequate.

9. COST OF REHEATING REFRIGERATION SYSTEM

The reheating refrigeration system does not require changes in other devices of the refrigeration system and the air conditioning parameters, thus it is not bearing the much costs except the cost of reheating system.

The cost for making the reheating system is combination of the cost of heat exchanger, reheater, conductor and insulation. These are having the cost of mainly its material. For higher heat transfer rate, materials having higher thermal conductivity are recommended. Here is list of some materials:

Material	Thermal conductivity (W/mK)
Aluminum	225
Brass	107
Copper	385
Cast iron	55-65
Steel	20-45
Silver	410
Urethane, rigid foam	0.026
Glass fiber	0.043
Soft rubber	0.13
Wood	0.17
Glass	0.78

The total cost for Air Conditioner having Reheating Refrigeration System may be about = Cost of Air conditioner without RRS + Cost of reheating device

Total cost of making the reheating device will be under Rs. 2000. Thus it will have effectiveness with efficiency as well as be economical.

11. NOMENCLATURE
COP = coefficient of performance

RRS = reheating refrigeration system

VCRS = vapour compression refrigeration system

h = enthalpy

 p-h diagram = pressure- enthalpy diagram

p = pressure

s = entropy

T = temperature

T-s diagram = temperature- entropy diagram

$ = American dollar

12. REFERENCES
1. Arora C.P., "Refrigeration and Air conditioning", 3rd edition, Tata McGraw-Hill Company, New Delhi.

2. Arora S.C. and Domkundwar, "Refrigeration and Air conditioning", Dhanpat Rai & Co. publication, New Delhi.

3. Yunus A. Cengel, "Heat and Mass Transfer", special Indian edition 2007, Tata McGraw-Hill Company, New Delhi.

4. Kumar D.S., "Heat and Mass Transfer", 7th revised edition, S.K. Kataria and sons publication, New Delhi.

5. Domkundwar and Domkundwar, "Refrigeration and Air conditioning Data Book (S.I. units)", Dhanpat Rai & Co. New Delhi.

6. Yadav R., "Thermodynamics & Heat Engines", volume 1, Central Publishing House, Allahabad.

7. Nag P.K., "Engineering Thermodynamics", Tata McGraw-Hill Company, New Delhi.

8. Khanna O.P., "Material Science", edition 2000, Dhanpat Rai & Co. publication, New Delhi.

A study on algorithms supported by CNG of Windows Operating System

Yasir Ahmad

Manav Bharti University, India

ABSTRACT

Cryptography Next Generation is an encryption Application Program Interface (API) which has been found out by Microsoft, in order to replace CryptoAPI in order to enable developers to add encoding, encryption and authentication to the windows based application that they develop. Cryptography Next Generation supports several algorithms in order to offer security to its users. This paper studies in detail the various algorithms supported by CNG of Windows Vista.

Keywords— Cryptography, Next Generation Cryprography, Windows Operating Systems.

1. CRYPTOGRAPHY NEXT GENERATION (CNG)

CNG was first introduced with Windows Vista. Cryptography Next Generation is extensible at several levels thus enhancing administrators to update, create and use custom cryptography algorithms in AD CS, IP Sec and SSL. One necessary function of cryptography is that it implements the United States government Suite B cryptographic algorithms including algorithms for digital signatures, hashing, key exchange and encryption. The United States NSA (National Security Agency) announced a combined set of asymmetric secret agreement and symmetric encryption also referred to as key exchange, hash and digital signatures functions for future United States government use referred to as Suite B. Suite B is used for the security of information configured as Top Secret and Secret and for private information that was configured Sensitive but uncategorized.

The APIs in Cryptography Next Generation can be used to do the following:

- Utilize and install extra cryptographic providers.
- Decrypt and encrypt data and create hashes.
- Store, create and retrieve cryptographic keys.

Cryptography Next Generation has the following capabilities among its several features:

- In the kernel model support for cryptography for use by boot processes, IP Sec and TLS/SSL. Cryptography Next Generation uses similar API in user and kernel mode for wholly supported features of cryptography. According to Microsoft, not all the functions of cryptography next generations can be referred from kernel mode.

- For organizations the Ability to use their own cryptographic algorithms or standard cryptographic algorithms implementation or to add new algorithms.
- With common criteria needs by storing and utilizing long lived keys in a protective process.
- Support for ECC (Elliptic Curve Cryptography) algorithms needed by United States government's Suite B.
- Support for CryptoAPI 1.0 algorithms and for TPM (Trusted Platform Module) computers which offers major isolation and storage in TPM.
- The capability to exchange the default random number generator by denoting specific random number Generator to use within chosen calls.
- Support for present algorithms supported by Crypto API.

Microsoft recommends that organizations do not deploy Suite B algorithms certificates before those organizations meet the following needs:

- Verify that any occurring PKI enhanced applications can use certificates that depend on cryptography next generation providers of cryptography.
- Verify that logon components of smart card can manage the algorithms of cryptography next generation.
- Before providing any certificates verify that occurring operating systems and CAs are capable to support Elliptic Curve Cryptography algorithms.

Presently organizations that do not have a PKI framework implemented can install a Windows Server 2008 CA once they assure that all occurring applications can support algorithms of Suite B. Organizations acquiring PKI with CAs on previous Windows Server operating systems must add a subordinate CA on a Windows Server 2008 computer. However they must continue using classic algorithms until their occurring CAs have been enhanced. One choice is to add a 2nd PKI and operate cross certification between the 2 CA hierarchies [1].

2. FEATURES OF CNG

Cryptography Next Generation has the following features that vary it from the legacy Crypto API [2].

2.1 Compliance and Certification

Cryptography Next Generation is aiming FIPS (Federal Information Processing Standards) 140-2 level 2 certification together with common criteria evaluation on chosen platforms. Other platforms will meet FIPS 140-2 level 1 certification.

2.2 Auditing

Cryptography Next Generation higher auditing capabilities to meet Common Criteria needs. The events are captured by the KSP (Key Service Provider) in user mode and includes:

- During operations of cryptography failures exist. The operations include decryption, signature verification, encryption, random number generation, encryption, hashing and key exchange.
- During keys testing errors exist. The tests include consistency, verification, parity checks and self tests.
- Destruction, generation, exporting and importing of key pairs.
- Writing and reading of persistent keys to and from the file system.

2.3 Kernel Mode Support

In kernel mode cryptography Next Generation supports cryptography. Kernel mode offers better performance for similar cryptographic features such as IP Sec and TSl/SSl. From kernel mode not all functions of cryptography next generation is being said.

2.4 Agility of cryptography

Cryptography next generation will support agility of cryptography or the capability to deploy new algorithms of cryptography for an occurring protocol such as TLS/SSL (Transport Layer Security/Secure Sockets Layer) or to disable algorithms if vulnerability is found with particular algorithm. This modification to operations of cryptography needs converting most standard cryptographic protocols such as Internet protocol security, S/MIME (Secure Multipurpose Internet Mail Extensions and Kerberos) to permit these protocols to take benefit of new algorithms possible in cryptography next generation.

2.5 Key Storage

For key storage cryptography Next Generation offers a model that supports both cryptography Next generation capable applications. The key storage router denotes the details for key access from both the used storage provider and the application.

3. NEW ALGORITHMS IN CNG

Cryptography Next Generation provides several newer algorithms most probably and notably most necessarily is support for Suite B. Some of the new algorithms of Cryptography Next Generation are:

3.1 RC2 Algorithm

RC2 is a symmetric block cipher configured by Ronald Rivest of RSA [3]. RSA configured RC2 as a direct exchange for DES by developing on the performance and offering a variable key size. RC2 is used commonly in S/MIME secure electronic mail and is referred to be 2 to 3 times as quick as DES. RC2 denotes to use the RC2 encryption algorithm that was developed by RSA Security. RC2 is a block cipher that encrypts data into 64 bits blocks. An encryption algorithm that separates down a message into blocks and encrypts every block is referred to as a block cipher. The key size of RC2 ranges from 8 to 256 bits. In SECURE/SAS a configurable size of key of 40 or 128 bits is used. The RC2 algorithm extends an individual message to a maximum of 8 bytes. RC2 is a proprietary algorithm developed by Data Security Inc. of RSA. RC2 encryption is an alternative to DES (Data Encryption Standard) encryption [4]. RC2 is vastly used algorithm that permits different key lengths but the security experts assume RC2 with little keys to be insecure.

3.2 RC4

RC4 is a stream cipher symmetric key algorithm. Stream ciphers exchange bytes or bits of plaintext streams into bytes or bits of cipher text. The stream cipher benefit is its speed in that only relies on the algorithm and not the performance of acquiring several plain texts. Similarly the stream cipher drawbacks are that it has less diffusion where a cryptanalyst can use techniques of language frequency distribution to separate it which can also lead to message fabrication. In 197 RC4 was developed by Ronald Rivest and kept as trade secret by Data Security of RSA. The below figure shows the RC4 algorithm:
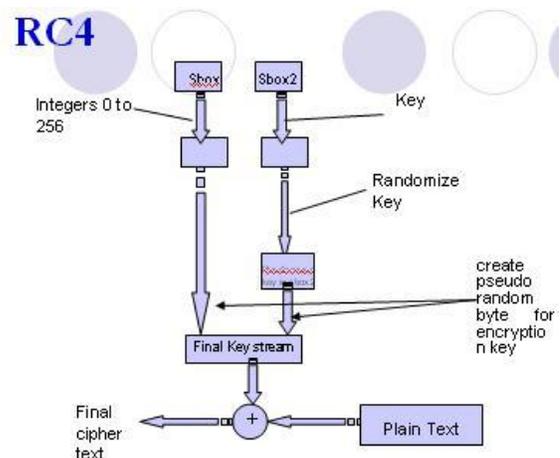

Fig 1: RC4 Algorithm

The RC4 algorithm operates in 2 phases such as encryption and key setup. In the 1st phase the RC4 algorithm uses a variable length key from 1 to 256 bytes to initialize a 256 byte state table Sbox (255). Then the 256 byte array is shuttled by N-number of mixing operations. Thus the key of

RC4 is limited to 40 bits and sometimes used as a 128 bits key. However it has the capability of using keys between 1 and 2048 bits. RC4 is used in several commercial packages of software such as Oracle SQL and Lotus Notes. In the 2nd phase the state table Sbox () is used for subsequent production of pseudo random bytes to produce a pseudo random stream which is XORed with the plain text to give the cipher text.

3.3 Advanced Encryption Standard

The AES (Advanced Encryption Standard) is the recent data security standard referred to as Federal Information Processing Standard 197 (FIPS 197) acquired worldwide by several private and public sectors for protective needs of data storage and secure data communications [5]. The Advanced Encryption Standard is used in several numbers of applications from mobile consumer products to high end users.

In 2001, NIST standardizes the symmetric key algorithm of Advanced Encryption Standard algorithm. The Advanced Encryption Standard denotes the Rijndael algorithm that can access 128 bits of data blocks using keys of 128-, 192- or 256 bit length. The Advanced Encryption Standard encipher exchanges data to an unintelligible form using the cipher key and the Advanced Encryption Standard decipher exchanges the cipher text back to plain text using similar cipher key. In Advanced Encryption Standard (AES) similar key is used for both decryption and encryption. Advanced Encryption Standard decryption and encryption are concerned on 4 various transformations applied again and again in a specific input data consequences and the data flows of decryption and encryption are not similar. The Advanced Encryption Standard also denotes an expansion key module to distribute keys for several iterations of the AES algorithm. The number of iterations of the Advanced Encryption Standard algorithm will vary depending on the input key length [6].   The below figure shows the AES diagram:
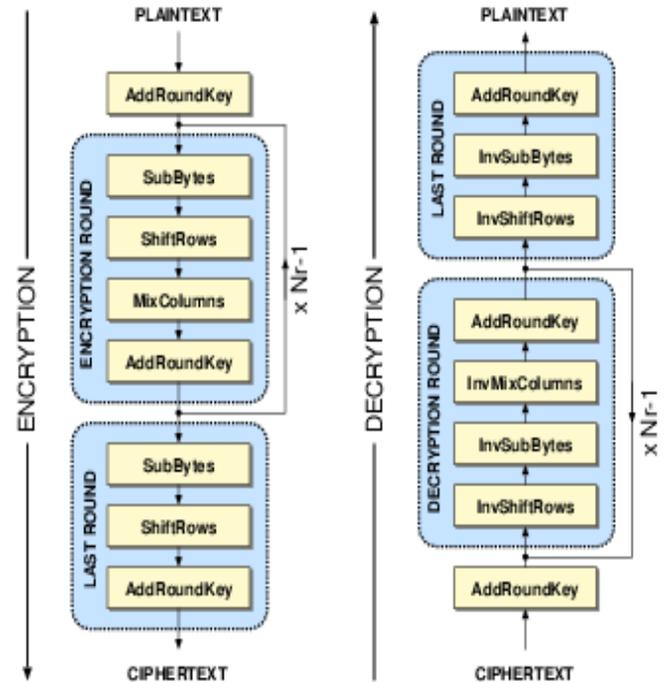

Fig 2: AES

Normally the Advanced Encryption Standard has 3 sizes of blocks such as AES-128, AES-192 and AES-256 bits [7]. From actual data to encrypted data the entire process consists 1 initial round,r-1 standard rounds and 1 final round. The major transformations consist of the following sections:

- Shift Rows: The input bytes are organized into 4 rows. Then according to its row value every row is rotated with a predefined step.
- Sub Bytes: By using a special design substitution box S (box) an input block is transformed byte by byte.
- Add Round Keys: The input block is XORed with the key is that round.
- Mix Columns: By using polynomial multiplication over GF (28) per column basis the 4 row structure which is organized is then transformed.

In the beginning operation there is one round Add Round Key operation and the standard round involves all 4 operations above. And the Mix Columns operation is removed in the final round operation while the other 3 operations remain. On the other hand for decryption the inverse transformations are applied. For quick implementation the round transformation can be parallelized. The entire block encryption is divided into various rounds. The design supporting AES-128 standard consists of 10 rounds.

3.4 Data Encryption Standard (DES)

Data Encryption Standard is the most known block cipher symmetric key which is recognized worldwide and it sets anterior in the middle 1970sas the first modern algorithm based on commercial grade with wholly and openly specified details of implementation [8]. It is defined by the FIPS 46-2 American Standard. The Data Encryption Standard is similar to 2 general concepts such as Feistel ciphers and product ciphers. Each cipher consists of iterating operation rounds or similar consequences operations. The product cipher's basic idea is to construct a composite encryption function by composing many easy operations that provides consequently but singly insecure protection. Basic operations consist of translations, linear transformations and transpositions, simple substitutions and modular multiplication.

A substitution permutation network is a product cipher consisting of several steps each consisting permutations and substitutions.

In a manner more than 2 transformations is integrated by product cipher enhancing that the out coming cipher is more protective than individual elements.

A block cipher consisting of internal function sequential repetition is an integrated block referred to as round function. The parameters consists of the block bit size n, r number of rounds, the Input key K bit size k from which r sub keys Ki from which r sub keys Ki are derived. The below figure shows the DES Algorithm:


Fig 3: DES Block Diagram

Data Encryption Standard needs a secret key that is 64 bits big but only 56 of those bits are actual key bits the remaining 8 bits are parity bits that assures the internal consistency of every byte of the key [9]. The Data Encryption Standard algorithm involves 16 rounds each one of which uses a varied 48 bit key to work its wonders. The actual 56 bit key is transformed into sixteen 48 bit keys as follows:

- The 56 bits of the key are permuted resulting in two 28 bit values the right hand key source and left hand key source.
- To acquire the key for every round the right hand key source and the left hand key source are shifted circularly each to one or two bits gaining a new right hand and left hand key source. The present round's key is acquired by operating a permutation on the integration of the present left hand key source and the present right hand key source gaining a 48 bit round key.

Data Encryption Standard is a complex algorithm. After the key for every round has been computed the real matter initiates. For CBC Mode every block is XOR'd with the past block's cipher text or for the 1st block with the IV. The XOR'd output is permuted and then categorized into a right hand half and a left hand half. The round key and the right hand half are used as the inputs to a complex numeric evaluation and its result is then XOR'd with the left hand half. The XOR'd output becomes the new right hand half, the previous right hand half becomes the new right hand half and the next round initiates. After these 16 rounds the final right hand and left hand half are concatenated and swapped, once gain permuted and the cipher text occurs. The above described steps are shown in the below figure:


Fig 4: DES Algorithm

3.5 MD2, MD4 and MD5 Hash Functions

MD2, MD4 and MD5 are one way hash functions which were established by Ron Rivest of RSA Data Security Inc. The MD stands for Message Digest [10]. These Algorithms are described in many internet RFCs which also consist of source

code of C language.

In 1989 Message Digest 2 hash algorithm was developed by Ronald Rivest to offer a secure hash function for 8 bit processors. MD2 pads the message so that its length is a multiple of 16 bytes. Then it computes a 16 byte checksum and adds it to the end of the message. Then the 128 bit message digest is produced by using the whole actual message along with the appended checksum. Cryptanalysts Attacks occurs against the MD2 algorithm.

Ronald Rivest developed Message Digest 4 algorithm in 1990 to support 32 bit processors and increase the security level. This developed algorithm is referred to as MD4. It first enhanced the message to assure that the message length is 64 bits smaller than a multiple of 512 bits. Then the MD4 algorithm processes 51 bit message blocks in 3 rounds of computation. The final output is a 128 bit message digest. Many mathematicians have published document papers errors in the full version of MD4 as well as inappropriate Message Digest 4 versions.

Similarly Ronald Rivest established the next version of the Message Digest algorithm in 1991 referred to as MD5. It also processes 512 bit message blocks but it uses 4 varied computation rounds to generate a digest to similar length as the MD4 and MD2 algorithms (128 bits). Message Digest 5 has similar needs of padding as Message Digest 4 and the length of the message must be 64 bits less than a multiple of 512 bits.

Message Digest 4 implements extra features of security that lowers down the message digest production speed essentially. Presently the cryptanalysts attacks describes that the Message Digest 5 protocol is concerned to collisions making it not a one way function [11].

3.6 Secure Hash Algorithm

SHA-1 Secure Hash Algorithm is another hash function which is also referred to as SHS (Secure Hash Standards) which was established by National Security Agency and National Institute of Standards and Technology and is used in government process of United States [12]. It can generate a 160 bit hash value from an arbitrary string length. Secure Hash Algorithm is structurally common to MD5 and MD4. Although it is about 25% slower than MD5 it is much more protective. It generates messages digests that are 25% higher than those produced by the Message Digest Functions making it more protective against attacks than Message Digest 5[13].

4. CONCLUSION

Cryptography is the science of Information security. The algorithms of cryptography are one of the main components in offering the mechanism of computer security. The cryptographic next generation offers exchange for the old API cryptography. BCrypt and NCryp are the sub divisions of cryptography next generation which offers low level

primitives of cryptography. The next generation cryptography offered by Windows makes use of several cryptographic algorithms in order to: 1) accomplish key management through quantum cryptography; 2) ensure that Elliptic Curve usage is as quicker and more portable than present asymmetric algorithms; 3) To establish cross web certification quicker several suite are viewed and; 4) finally move word to word suits of standardization to handle the implementation quality of the cryptographic algorithm in electronic mail at various security levels and authentications of cross web.

REFERENCES

1) Price J A, Price B and Fenstermacher S, *Mastering Active Directory for Windows Server 2008, John Wiley & Sons, New York, p 11-40, (2008).*
2) Komar B, *Windows Server® 2008 PKI and Certificate Security, O'Reily Media Inc, USA, (2008).*
3) Burnett M and Foster J C, *Hacking the code: ASP.NET web application security, Syngress, USA, p 165, (2004).*
4) SAS, *SAS 9.2 Language Interfaces to Metadata, SAS Institute Inc., USA, p 31, (2009).*
5) Saadi L, *Stealth Ciphers, Trafford Publishing, Canada, (2004).*
6) Malepati H, *Digital Media Processing: DSP Algorithms Using C, Newns, USA, p 37,(2010).*
7) Flynn M J and Luk W, *Computer System Design: System-on-Chip, John Wiley & Sons, New Jersey, p 251, (2011).*
8) Patel D R, *Information Security: Theory And Practice, Prentice Hall, New Delhi, p 47, (2008).*
9) Frankel S, *Demystifying the IPsec puzzle, Artech House Inc., USA, p 70, (2001).*
10) Hughers L J, *Actually useful Internet security techniques, New Riders Publishing, USA, p 55, (1995).*
11) Stewart J M, Tittle E and Chapple M, *CISSP: Certified Information Systems Security Professional Study Guide, John Wiley Publishing, USA, p 416, (2011).*
12) Stanger J M, Lane P T and Crothers T, CIW: *security professional: study guide, John Wiley & Sons, USA, p 71, (2002).*
13) Singh A and Singh B, *Identifying Malicious Code Through Reverse Engineering, Springer, USA, p 68, (2009).*

Consequence of Corrosion on Dynamic Behaviour of Steel Bridge Members

J.M.R.S. Appuhamy[1], M. Ohga[1], P. Chun[1] and P.B.R. Dissanayake[2]

[1]Department of Civil & Environmental Engineering, Ehime University,
Bunkyo-cho 3, Matsuyama 790-8577, Japan
[2]Department of Civil Engineering, University of Peradeniya,
Peradeniya 20400, Sri Lanka

ABSTRACT

Deterioration of steel bridge infrastructures constitutes a major worldwide problem in transportation engineering and maintenance management industry. Corrosion and fatigue cracking may be the two most important types of damage in aging structures. Various kinds of failures and the need of expensive replacements may occur even though the amount of metal destroyed is quite small. One of the major harmful effects of corrosion is the reduction of metal thickness leading to loss of mechanical strength and structural failure, causing severe disastrous and hazardous injuries to people.

Furthermore, some recent earthquakes demonstrated the potential seismic vulnerability of some types of steel bridges. Corrosion and its effects can trigger the damages caused by earthquakes and it would be vital to understand the behavior of existing steel bridges which are corroding for decades in future severe seismic events. Therefore, this paper proposes a methodology to estimate the reduction of seismic strength capacities due to corrosion, which can be used to make rational decisions about the maintenance management plan of steel highway and railway infrastructures.

Keywords – **Corrosion, Earthquakes, Seismic resistance, Steel bridges, Maintenance management.**

I. INTRODUCTION

As bridges are regarded as critical components of a transport network, the safety of bridges is crucial to people's daily life and to the national economy. Steel girder bridges, like other structures, deteriorate over time due to environmental effects, material fatigue and overloading. As a consequence, their closure or traffic capacity reduction causes major inconveniences for the users and result in significant losses to the economy. Furthermore, major steel bridges are rather costly to construct, maintain and rehabilitate. Therefore, the design, maintenance, management and rehabilitation of this kind of structures are complex and challenging tasks. Consequently, they have to be performed using state-of-the-art technical solutions and the best practice gained over many years of experience.

Exposure of a steel structure to the natural environment and inadequate maintenance will cause corrosion. Corrosion is one of the most important causes of deterioration of steel girder bridges which affects their long term mechanical performance, usability and durability [1, 2]. It has been pointed out that the corrosion can lead to cracking (fracture), yielding or buckling of members which can result in stress concentration, changes in geometric parameters and a build-up of the corrosion products. These parameters are critical for the member's ability to resist load effects [3]. Further, it is known that the corrosion wastage and stress concentration caused by the surface irregularity of the corroded steel plates influence the remaining strength of corroded steel plates [4]. Controlling corrosion on bridge structures can prevent premature failure and lengthen their useful service life, both of which save money and natural resources, and promote public safety. Therefore, careful evaluation of remaining load-carrying capacities of corroded steel structures is of high importance in transportation and maintenance engineering.

Efficient maintenance, repair and rehabilitation of existing bridges require the development of a methodology that allows for an accurate evaluation of the load carrying capacity and prediction of remaining life. In the past few decades, several experimental studies and detailed investigations of corroded surfaces were done by some researchers in order to introduce methods of estimating the remaining strength capacities of corroded steel plates [5 - 8]. Recent worldwide severe earthquakes have shown that steel bridges can be vulnerable. Especially during the Great Hanshin Earthquake of January 17, 1995, steel bridge structures suffered various kinds of damages which they never experienced before [9]. Furthermore, many collapses, superstructure failures and joint failures were observed in San Francisco due to the Loma Prieta Earthquake on October 17, 1989 [10]. Therefore it is evident that the failure risk associated with severe corrosion of steel bridge structure under mega earthquake events could not be underestimated.

Despite these uncertainties and variations, one can learn from past earthquake damage, because many types of damage occur repeatedly. By being aware of typical vulnerabilities that bridges have experienced, it is possible to gain insight into structural behavior and to identify potential weaknesses

in existing and new bridges. Historically, observed damage has provided the impetus for many improvements in earthquake engineering codes and practice. Although there is no published information on the seismic performance of severely corroded members, practicing engineers have apparently not questioned whether rusted members can still exhibit the dynamic behavior and hysteretic energy dissipation capacity commonly relied on for seismic resistance [11, 12]. Therefore this paper aims to investigate the effect of severe corrosion on the remaining seismic capacities of existing steel bridge infrastructures.

II. PRELIMINARY INVESTIGATION

A steel girder bridge of Ananai River in Kochi prefecture on the shoreline of the Pacific Ocean, which had been used for about hundred years, was used for this study. This bridge had simply supported steel plate girders with six spans, with each of 13.5 m. It was constructed as a railway bridge in 1900, and in 1975 changed to a pedestrian bridge, when the reinforced concrete slab was cast on main girders. The bridge was dismantled due to serious corrosion damage in year 2001.

2.1 Corroded Specimens for Tensile Test

The specimens for tensile loading tests were cut out from the portion adjacent to end support of a plate girder of Ananai River Bridge. There, 21 flange specimens (F1-F21) and 5 web specimens (W1-W5) were fabricated from the cover plate on upper flange and web plate respectively. Here, the flange and web specimens have the widths ranged from 70-80mm and 170-180mm respectively. Test specimen configuration is shown in Figure 1.
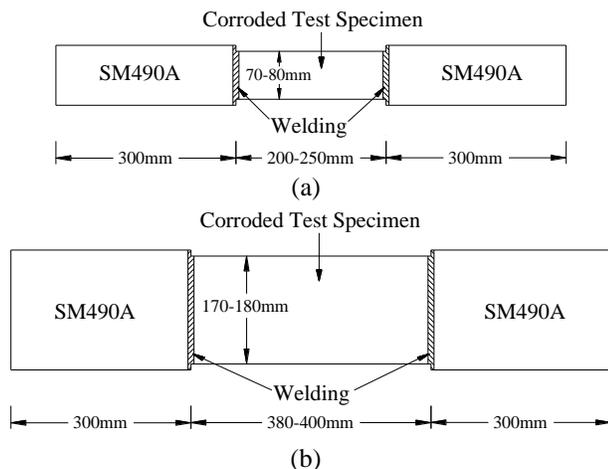

Fig. 1: Dimensions of (a) flange and (b) web test specimens

In addition, 4 corrosion-free specimens (JIS5 type) were made of each two from flange and web, and the tensile tests were carried out in order to clarify the material properties of test specimens. The material properties obtained from these tests are shown in Table 1.

Table 1: Material properties

Specimen	Elastic modulus /(GPa)	Poisson's ratio	Yield stress /(MPa)	Tensile strength /(MPa)	Elongation at breaking /(%)
Corrosion-free plate (flange)	187.8	0.271	281.6	431.3	40.19
Corrosion-free plate (web)	195.4	0.281	307.8	463.5	32.87
SS400 JIS	200.0	0.300	245~	400~510	–

2.2 Classification of Corrosion Levels

In this study, all specimens were categorized into typical 3 corrosion types concerning their corrosion conditions and minimum thickness ratio, μ (minimum thickness/ initial thickness). Figure 2 shows the relationship between the nominal ultimate stress ratio (σ_{bn}/σ_b) and the minimum thickness ratio (μ), where σ_{bn} is the nominal ultimate stress and σ_b is the ultimate stress of corrosion-free plate.
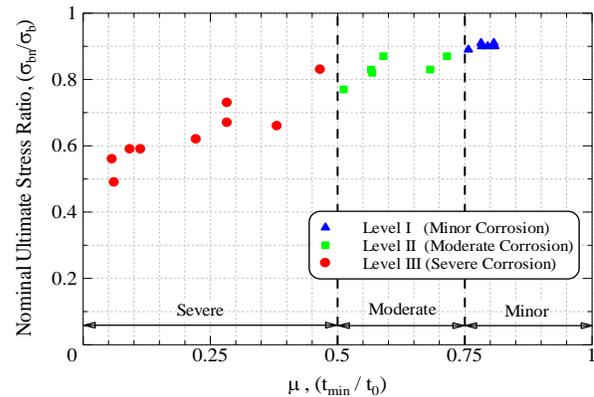

Fig. 2: Relationship of ultimate stress ratio & minimum thickness ratio (μ)

The corrosion conditions with the minimum thickness ratio, $\mu > 0.75$ are defined as 'minor corrosion'. And the 'moderate corrosion' type is defined when the minimum thickness ratio, $0.75 \geq \mu \geq 0.5$. Further, the 3rd corrosion type with the minimum thickness ratio, $\mu < 0.5$ is defined as 'severe corrosion' [8]. There, the initial thickness (t_0) of flange and web specimens are 10.5mm and 10.0mm.

2.3 Numerical Analysis

The non-linear finite element analyses were performed for the specimens with different corrosion conditions in order to clarify the yield and tensile strengths, failure surfaces and the ultimate behaviors of corroded members. The 3D isoparametric hexahedral solid element with eight nodal points (HX8M) and updated Lagrangian method based on incremental theory were adopted in these analyses. Non linear elastic-plastic material, Newton-Raphson flow rule and Von Mises yield criterion were assumed for material

properties. Further, an automatic incremental-iterative solution procedure was performed until they reached to the pre-defined termination limit.

The analytical models with different length (L) and width (W) dimensions were modeled with their respective corrosion conditions and 2mm regular mesh pattern was adopted for all analytical models. One edge of the member's translation in X, Y and Z directions were fixed and only the Y and Z direction translations of the other edge (loading edge) were fixed to simulate with the actual experimental condition. Then the uniform incremental displacements were applied to the loading edge. Yield stress σ_y = 294.7 [MPa], Elastic modulus E = 191.6 [GPa], Poisson's ratio ν = 0.276 were applied to all analytical models, respectively.

2.3.1 Ductile Fracture Criterion

The "Stress Modified Critical Strain Model (SMCS)" was proposed by Kavinde *et al.* [13], to evaluate the initiation of ductile fracture as a function of multiaxial plastic strains and stresses. This method was adopted in this analytical study. In SMCS criterion, the critical plastic strain ($\varepsilon_p^{critical}$) is determined by the following expression:

$$\varepsilon_P^{Critical} = \alpha \cdot Exp\left(-1.5\frac{\sigma_m}{\sigma_e}\right) \qquad (1)$$

Where, α is toughness index and the stress triaxiality T = (σ_m/σ_e), a ratio of the mean or hydrostatic stress (σ_m) and the effective or von Mises stress (σ_e). The toughness index α is a fundamental material property and hence obtained from the tensile test conducted for the non corroded specimen and obtained as follows:

$$\alpha = \frac{\varepsilon_P^{Critical}}{Exp\left(-1.5\frac{\sigma_m}{\sigma_e}\right)} \qquad (2)$$

The ultimate strength of each corroded specimen was calculated accordingly by using the SMCS criterion and compared with their experimental ultimate capacities to understand the feasibility of the numerical modeling approach for remaining strength estimation of corroded steel plates with different corrosion conditions.

2.3.2 Analytical Results

First, analytical modeling of the non-corroded specimen was done with above described modeling and analytical features to understand the accuracy of the adopted procedure. It was found that the analytical model results were almost same as the experimental results with having a negligible percentage error of 0.03% and 0.02% in yield and tensile strengths respectively. Then, all other experimentally successful specimens were modeled accordingly and their yield and ultimate strengths were compared with the experimentally obtained values.

Figure 3 shows the comparison of experimental and analytical the load-displacement curves of three specimens F-14, F-13 and F-19 with minor, moderate and severe corrosion conditions respectively. There, it can be seen that a very good agreement of experimental and analytical load-displacement behaviors for all 3 classified corrosion types can be obtained. Here, the percentage errors in yield and tensile strength predictions of the analytical models of three corrosion types are 0.53% and 0.03% in F-14, 2.96% and 0.70% in F-13 and 3.20% and 5.53% in F-19 respectively. Therefore, it is revealed that this analytical method is accurate enough and hence can be used to predict the yield, tensile and dynamic behaviors of actual corroded specimens more precisely.
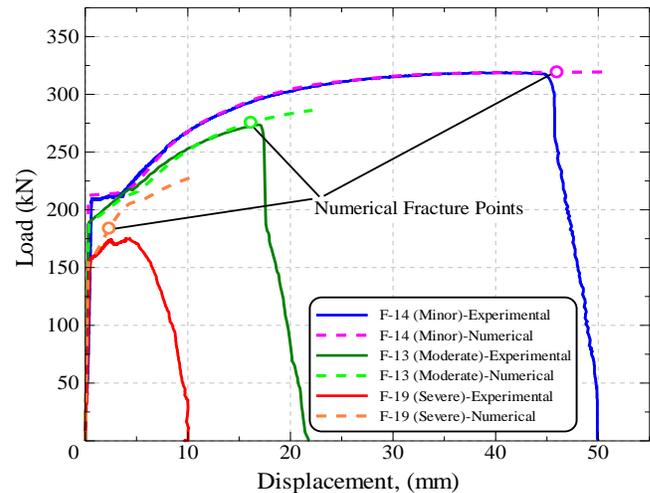

Fig. 3: Comparison of experimental and analytical load-displacement curves

2.4 Modeling of Corrosion Conditions

Two corrosion condition modeling (CCM) parameters were defined to model the corroded surface considering the material loss due to corrosion and stress concentration effect [14]. Figure 4(a) shows the variation of diameter of the maximum corroded pit (D) vs. maximum corroded depth ($t_{c,max}$) and Figure 4(b) shows the normalized average thickness (t_{avg}/t_0) vs. normalized maximum corroded depth ($t_{c,max}/t_0$). Both figures show a very good linear relationship and hence these parameters were used to develop an analytical model which can be used to predict the yield, ultimate and dynamic behaviors of corroded steel plates with different corrosion conditions.

Therefore, by considering the Figures 4(a) and (b), the two equations for the corrosion condition modeling (CCM) parameters can be defined as:

$$D^* = 5.2\ t_{c,max} \qquad (3)$$

$$t^*_{avg} = t_0 - 0.2\ t_{c,max} \qquad (4)$$
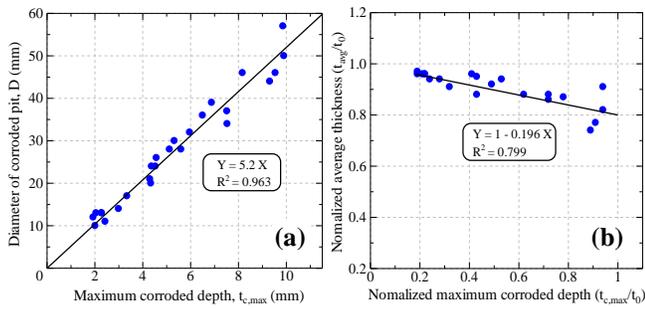
Fig. 4: Relationship of (a) D vs. $t_{c,max}$ and (b) normalized t_{avg} vs. $t_{c,max}$

Where D^* and t^*_{avg} are the representative diameter of maximum corroded pit and representative average thickness respectively.

Figure 5 shows the analytical model, which is developed with the above CCM parameters with different corrosion conditions. Here, the maximum corroded pit was modeled with the representative diameter (D^*) which could account the stress concentration effect and the material loss due to corrosion was considered by using the representative average thickness parameter (t^*_{avg}). The same modeling features and analytical procedure as described in section 2.3 were adopted for the analyses.
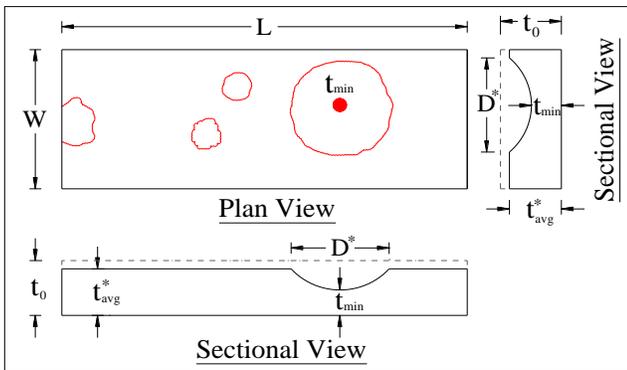
Fig. 5: Analytical model with CCM parameters

Figure 6 shows the comparison of load-displacement behavior of experimental and analytical results with proposed model for 3 members F-14, F-13 and F-19 with minor, moderate and severe corrosion conditions respectively. It was revealed that a very good comparison of the load-displacement behavior can be seen for the all three classified corrosion types for the proposed analytical model too. Here, the percentage errors in yield and tensile strength predictions of the proposed analytical model for the three corrosion types are 0.13% and 0.83% in F-14, 0.38% and 1.01% in F-13 and 3.51% and 2.69% in F-19 respectively.
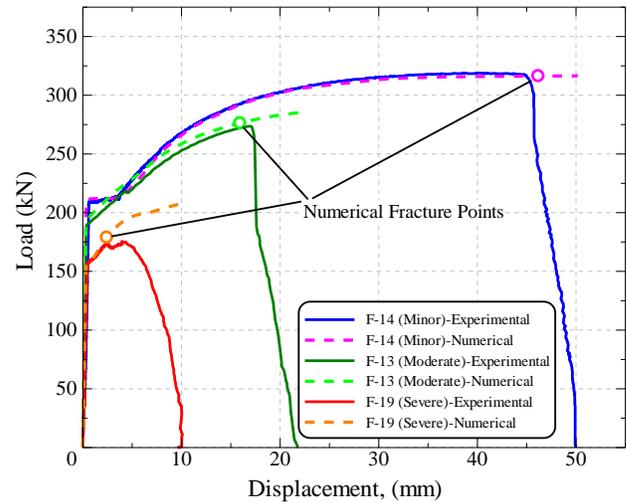

Fig. 6: Comparison of experimental and analytical load-displacement curves with proposed analytical model with CCM parameters

Then, all experimentally successful specimens were modeled accordingly with the proposed analytical model and their yield and ultimate strengths and failure surfaces were compared with the experimental results. Figure 7 shows the comparison of ultimate load capacities of all specimens in experimental and numerical analyses with proposed model with CCM parameters. Here, having a coefficient of correlation of $R^2 = 0.992$ indicate the accuracy of the proposed model and the possibility of the use of proposed analytical model instead of the model with detailed corroded surface measurements.
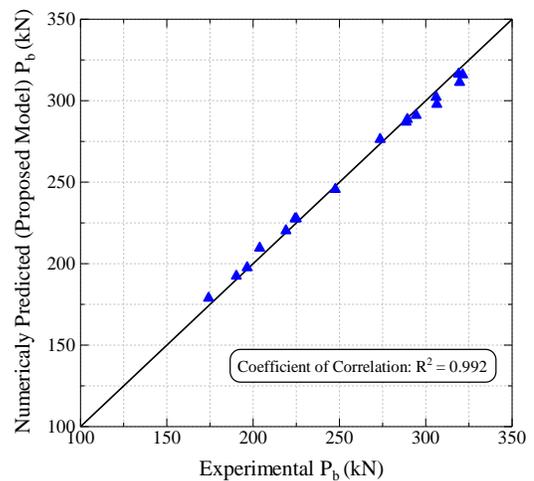

Fig. 7: Comparison of experimental and analytical ultimate load capacities with proposed analytical model with CCM parameters

III. DYNAMIC ANALYSIS

The seismic capacity of existing bridges is important after an earthquake both to provide emergency supplies and transportation to society as well as to ensure structural safety in engineering terms. Better seismic capacity of bridges results in less structural damage and reduced impacts following an earthquake disaster. Therefore, an evaluation of the seismic capacity for numerous existing bridge inventories can be regarded as a fundamental basis for mitigating future losses caused by earthquakes.

3.1 Seismic Analysis with Historical Earthquakes

In this study, three major historical earthquake records were used to understand the dynamic response of Ananai River Bridge in Kochi prefecture, Japan. The seismic analyses were performed considering the earthquake excitations caused by; 1995 Kobe earthquake (M_w=6.9), 1989 Loma Prieta earthquake (M_w=6.9) and 1940 El-Centro earthquake (M_w=7.1). Figure 8 shows the acceleration histories of 1995 Kobe earthquake used in this analysis.
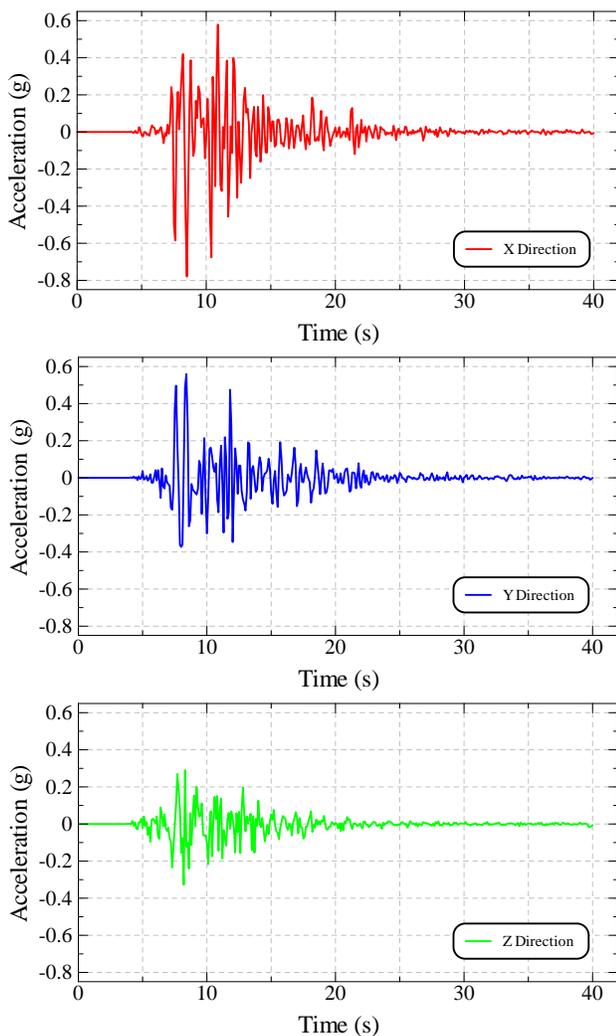
The seismic response analysis is performed in two distinct stages. A natural frequency analysis is performed first. This is used to calculate the first 100 (or more; until more than 90% of total mass participation occurs) natural modes of vibration of the structure. The eigenvalues (frequencies) and eigenvectors (mode shapes) are stored and used in the subsequent IMDPlus analysis. The second phase of the analysis utilizes the IMDPlus option which performs enhanced time domain solutions using Interactive Modal Dynamics (IMD). In the IMDPlus solution, the structure is subjected to a support condition excitation governed by time histories of acceleration in the model global axes. In this example, the seismic excitation is applied directly to the bases of the structure using the first 40 seconds of each earthquake.

3.2 Results of Primary Seismic Analysis

The primary seismic analysis was performed with the three historical earthquakes as described above and the behavior of the structure was studied. In each analysis, critical members were identified considering the whole excitations of the corresponding earthquake records. It was noticed that, even though the main steel girders are well behaved due to earthquake loading, cross girder members were deformed in transverse direction. One example of the ultimate stress distribution of the bridge structure and the displacement history obtained for the critical member for the seismic analysis of 1995 Kobe earthquake is shown in Figure 9.


Fig. 8: Acceleration histories of 1995 Kobe earthquake
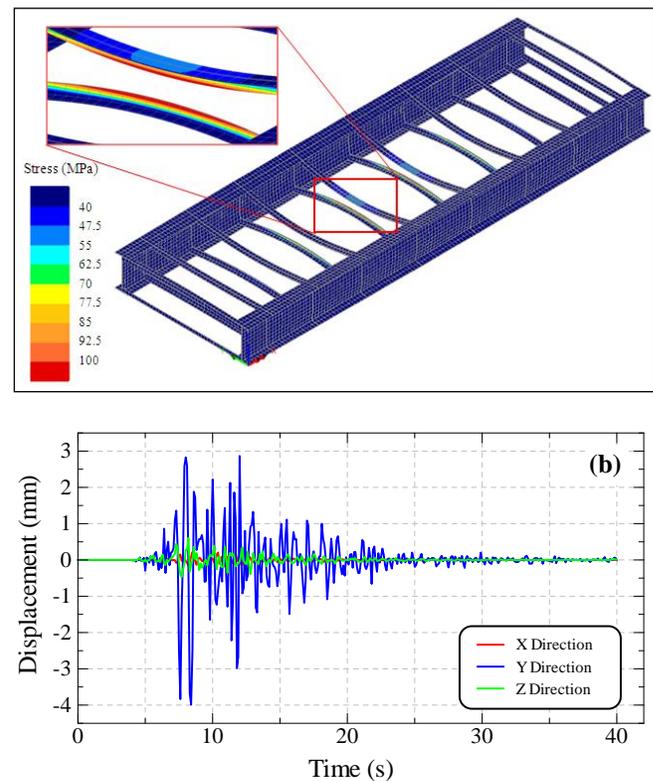
Fig. 9: (a) Ultimate stress distribution of steel girder and (b) displacement histories of the critical member [Kobe eqk.]

3.3 Secondary Seismic Analysis

3.3.1 Analytical Models

Five different analytical models were developed and analyzed with the same modeling features, analytical procedure, in order to understand the seismic strength deterioration with the severity of corrosion condition. Here, the seismic model SM-1 with $\mu=1$ represents the model without corrosion and used as the standard model in this analysis to compare with other SM models. The different corrosion conditions were adopted by using different minimum thickness ratio (μ) values and the CCM parameters; representative diameter (D^*) and representative average thickness parameter (t^*_{avg}), were used to model the stress concentraion effect and material loss due to corrosion as decribed in section 2.4. The details of those analytical models which are considered for this parametric study are shown in Table 2. The initial thickness (t_0) of the different sismic models (SM) was considered as 10.5mm. Here, the maximum corroded depth ($t_{c,max}$) can be expressed as follows:

$$\mu = \frac{t_{min}}{t_0} = \frac{(t_0 - t_{c,\,max})}{t_0}$$

$$t_{c,\,max} = t_0(1 - \mu) \qquad (5)$$

Table 2: Details of analytical models

Model	μ (t_{min}/t_0)	$t_{c,max}$ (mm)	D^* (mm)	t^*_{avg} (mm)
SM-1	1.0	0.0	0.0	10.5
SM-2	0.75	2.63	13.65	9.98
SM-3	0.50	5.25	27.30	9.45
SM-4	0.25	7.88	40.95	8.93
SM-5	0.05	9.98	51.87	8.51

3.3.2 Analytical Results

The secondary seismic analysis for each seismic model (SM-1~SM-5) with different corrosion conditions, was performed by using the primary displacement histories obtained for each historical earthquake. The analyses were conducted until they reached to their pre-defined termination limits and the load-displacement behavior for each model was obtained.

Figure 10 shows the load-displacement curves for; (a) Kobe earthquake, (b) Loma Prieta earthquake and (c) El-Centro earthquake for each seismic model. They clearly show that the dynamic behavior of each model was affected by different corrosion conditions. Further, it was noted that their residual strengths were significantly affected by the different levels of corrosion conditions attributed to each SM model.
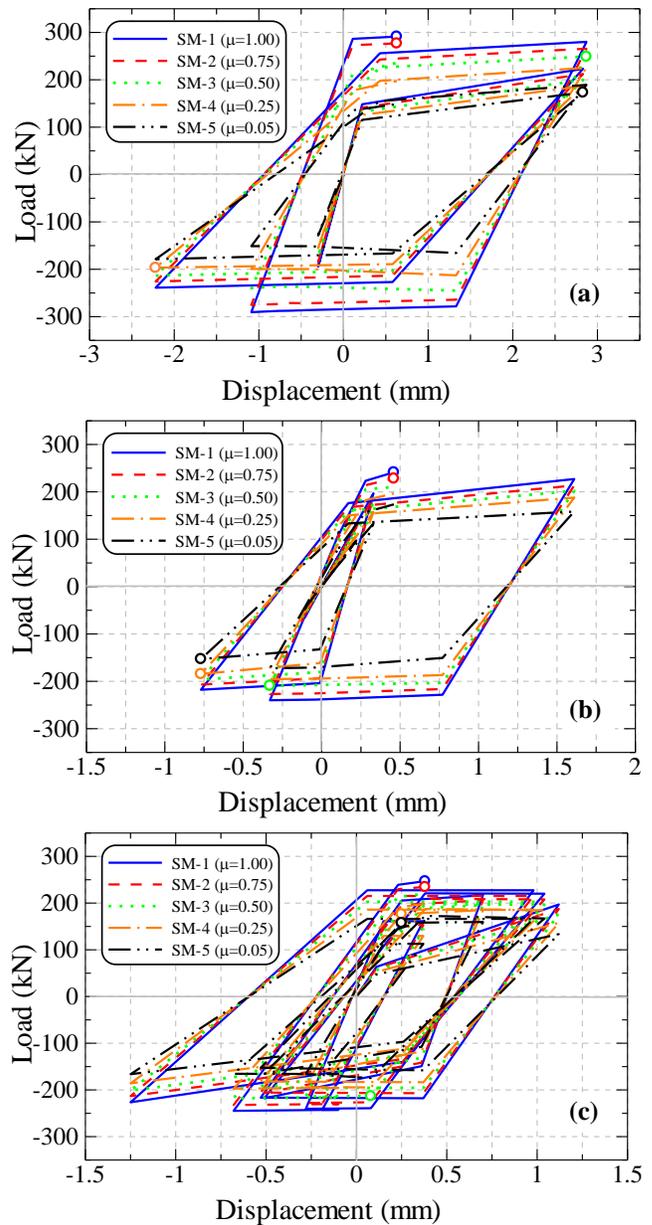

Fig. 10: Comparison of load-displacement curves of different seismic models for (a) Kobe earthquake, (b) Loma Prieta earthquake and (c) El-Centro earthquake

3.4 Discussion

The percentage seismic strength reductions (SSR%) in yield and ultimate strength states of each SM model are shown in Figure 11(a) and 11(b) respectively. There, the percentage strength reduction results of different seismic models and analytical results of tensile specimens developed with CCM parameters are shown. It was noted that percentage strength reductions (%SR) of tensile analysis have linear relationship in both yield and ultimate strength estimations. But it was noted that the %SSR in yield and ultimate states are non-linearly increased with increase of corrosion levels.
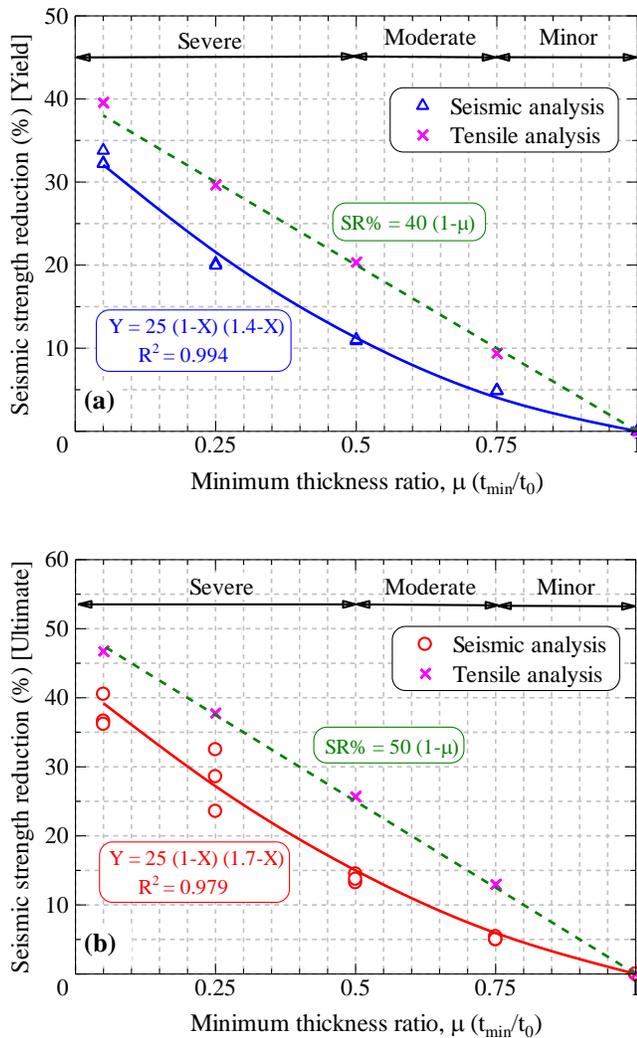
Fig. 11: Relationship of percentage (a) yield; (b) ultimate, seismic strength reduction vs. minimum thickness ratio (μ)

The remaining yield and ultimate seismic strength capacities are decreased with the severity of corrosion condition for all three different earthquake loading histories. Furthermore, these results divulged that having a corrosion pit of minimum thickness ratio $\mu=0.05$, could reduce up to 30% ~ 40% of its' original yield and ultimate strengths respectively. It is noted from Figure 11 that two very good relationships can be obtained for remaining yield and ultimate seismic strength capacities with high accuracy. Therefore, by considering Figure 11(a) and 11(b), two equations for estimating remaining yield and ultimate seismic strength capacities can be derived as:

$$\%SSR_{yield} = 25\ (1-\mu)\ (1.4-\mu) \qquad (6)$$

$$\%SSR_{ultimate} = 25\ (1-\mu)\ (1.7-\mu) \qquad (7)$$

As the proposed strength reduction equations only requires the measurement of minimum thickness ratio, μ (minimum thickness/ initial thickness), which is an easily measurable parameter through a quick and careful site investigation, this method can be used as a simple and reliable method to predict the yield and ultimate seismic behaviors of corroded steel members more easily and precisely. Furthermore, the seismic strength reduction charts give a good indication about the % strength reduction according to the severity of corrosion and hence bridge engineers would be able to quickly decide the required maintenance works of corroded members, in order to assure the adequate safety of the infrastructure.

IV. CONCLUSIONS

The corroded surface measurements and tensile tests were conducted on many wide specimens with different corrosion conditions, which are obtained from a plate girder which had been used for about 100 years with severe corrosion. Non-linear FEM analysis was carried out to understand the mechanical behavior, stress distribution, ultimate behavior etc. for members with different corroded conditions. The yield and ultimate behaviors of steel bridge members with different corrosion conditions were studied under different earthquake loadings. The main conclusions obtained from this study can be summarized as follows.

1. A very good agreement between experimental and non linear FEM results can be seen for all three classified corrosion types. So, the adopted numerical modeling technique can be used to predict the remaining strength capacities of actual corroded members accurately.

2. Even though, the steel main girders behave well during severe earthquakes, some types of transverse members could become critical and hence could suffer various structural damages.

3. Furthermore, the corrosion and its stress concentration effect will trigger those damages significant and could even lead to the collapse of total structure.

4. The percentage reduction of yield and ultimate strength capacities due to corrosion, under earthquake loadings can be defined as:

$$\%SSR_{yield} = 25\ (1-\mu)\ (1.4-\mu)$$

$$\%SSR_{ultimate} = 25\ (1-\mu)\ (1.7-\mu)$$

As the above equations have only a single variable, minimum thickness ratio (μ), which is an easily measurable parameter, it will reduce the contribution of the errors occurred during the practical investigation of a corroded member. Further this method is simple and hence can be used for the maintenance management of steel bridge infrastructures with better accuracy.

REFERENCES

[1] K. Fujii, T. Kaita, H. Nakamura and M. Okumura, A model generating surface irregularities of corroded steel plate for analysis of remaining strength in bridge maintenance, *Proceedings of EASEC-9*, Indonesia, 9, 2003, 32-38.

[2] R. Rahgozar, Remaining capacity assessment of corrosion damaged beams using minimum curves, *Journal of Constructional Steel Research*, 65, 2009, 299-307.

[3] Y. Sharifi and R. Rahgozar, Remaining moment capacity of corroded steel beams, *International Journal of Steel Structures*, 10(2), 2010, 165-176.

[4] A. Kariya, K. Tagaya, T. Kaita and K. Fujii, Mechanical properties of corroded steel plate under tensile force, *Proceedings of the 3rd International Structural Engineering and Construction Conference*, Japan, 2005, 105-110.

[5] M. Matsumoto, Y. Shirai, I. Nakamura and N. Shiraishi, A proposal of effective thickness estimation method of corroded steel member, *Bridge and Foundation Engineering*, 23(12), 1989, 19-25. (in Japanese)

[6] A. Muranaka, O. Minata and K. Fujii, Estimation of residual strength and surface irregularity of the corroded steel plates, *Journal of Structural Engineering*, 44A, 1998, 1063-1071. (in Japanese)

[7] A. Kariya, K. Tagaya, T. Kaita and K. Fujii, Basic study on effective thickness of corroded steel plate and material property, *Annual Conference of JSCE*, Japan, 2003, 967-968. (in Japanese)

[8] J.M.R.S. Appuhamy, T. Kaita, M. Ohga and K. Fujii, Prediction of residual strength of corroded tensile steel plates, *International Journal of Steel Structures*, 11(1), 2011, 65-79.

[9] C. Miki and E. Sasaki, Fracture in steel bridge piers due to earthquakes, *International Journal of Steel Structures*, 5(2), 2005, 133-140.

[10] J.P. Moehle and M.O. Eberhard, Earthquake damage to bridges, *Bridge Engineering Handbook,* CRC Press, 2000, 1-34.

[11] M. Bruneau and S.M. Zahrai, Effect of severe corrosion on cyclic ductility of steel, *Journal of Structural Engineering*, 1997, 1478-1486.

[12] S.M. Zahrai, Cyclic strength and ductility of rusted steel members, *Asian Journal of Civil Engineering*, 4(2-4), 2003, 135-148.

[13] A.M. Kavinde and G.G. Deierlein, Void growth model and stress modified critical strain model to predict ductile fracture in structural steels, *Journal of Structural Engineering*, 132(12), 2006, 1907-1918.

[14] J.M.R.S. Appuhamy, M. Ohga, T. Kaita and P.B.R. Dissanayake, Reduction of ultimate strength due to corrosion - A finite element computational method, *International Journal of Engineering*, 5(2), 2011, 194-207.

A MODEL TO SUBJUGATE TCP STARVATION FOR WIRELESS MESH NETWORKS

Sk.MANSOOR RAHAMAN[1] , B.RAMESH BABU[2], A.VENKAIAH NAIDU[3], K.RAJASEKHARA RAO[4]

[1]Student, Department Of CSE, K.L.university
[2]Assoc.Prof, Department Of CSE, K.L.university
[3]Student, Department Of CSE, K.L.university
[4]Principal, K.L.university

ABSTRACT

wireless mesh networks is an emerging technology in wireless communications, it provides high speed internet access at low cost. In basic topology of wireless mesh network the node which is two hop away from the gate way will starves compared to the nodes which are in direct contact with gate way node. The nodes which are one hop away from gateway node will have high through put compare to the nodes which are two hops away. So in this bandwidth is utilized by all the nodes efficiently so we are proposing an algorithm known as fair binary exponential backoff algorithm, to reduce starvation and utilization of bandwidth efficiently by all the nodes irrespective of their position and hop distance from the gateway.

KEYWORDS: *FBEB, MAC, WIRELESS MESH NETWORKS, TOPOLOGY, GATEWAY*

I.   INTRODUCTION:

Coordination of medium access:

This is the key component of a MAC protocol, which involves many different tasks depending on what type of MAC protocols need to be designed. for a reservation based MAC protocol, the key task is to assign resources such as codes, time slots, subcarriers or channels, to users such that the network throughput is maximized, but their QOS is also satisfied. To this end, many other algorithms in the physical layer need to be considered, for example, power control, adaptive coding and modulation, etc. in addition, functions in the network and transport layers also need to be considered. for example, a TDMA MAC may impact slow start performance of TCP owing to the significant differences of round trip time (rtt) before and after resource allocation these demands imply that cross-layer design between MAC and other protocol

Layers are important.[12][1]

II.   Random Access MAC:

for a random access MAC protocol such as csma/ca, the key issue is to find out the best solution for

minimizing collision and fast recovery from collision in case it still happens. Since no reservation is available, collision becomes severe when the number of users increases, and thus significantly degrades the throughput performance. No QOS can be guaranteed. However, random access MAC protocols have two main advantages. First is their simplicity. No separate signaling and reservation schemes are needed in the protocol. Second is the compatibility with connectionless (datagram) networks such as the internet.[3]

On the contrary, a reservation-based MAC protocol always has the problem of how to do integration with a connectionless network. for example, if a TDMA MAC is used, whenever a tcp session starts, it has to wait for the allocation to be done. such a delay is not tcp-friendly, because tcp assumes that the network is congested before even resource allocation is completed. Another example is when video traffic is sent through a TDMA MAC into the internet; the MAC has no way of knowing its bandwidth and QOS requirements. Without such information, reservation cannot be correctly done, unless adaptive resource

Estimation and dynamic time slot allocation are designed interactively. however, a random access MAC protocol does not have any of these problems, because a packet starts its transmission process as it arrives.[4]

Network formation and association. This component is actually the network management part for a MAC protocol. it takes care of network formation and association/

Disassociation of a node to/from the network when a node joins/leaves the network. This is particularly important for WMNS. Without network formation and association, network nodes cannot recognize each other and accordingly start their MAC protocol. A MAC protocol can be implemented in two types of architecture. In the classical implementation architecture, a MAC protocol is implemented in software (MAC driver), firmware, and hardware. Usually, packet queuing, network formation, node association, and so on, are done in the driver. Timing critical functions, e.g., time slot generation, back off procedures, etc., are performed in the firmware. The actual real-time operation of the MAC protocol is done in the

hardware. for example, when a back off counter is determined, the exact decrement of this counter is done in the hardware in order to achieve high accuracy. thus far, many companies have tried to pull more functions in the firmware into the driver level so that the driver has more freedom to control/modify the MAC protocol. this type ofme thod is usually called a "soft MAC" implementation. However, since many key functions are still located in firmware, the timing critical part of the MAC protocol is hard to modify. owing to the mesh networking topology, the design of a MAC protocol for WMNS is more challenging than that for a single-hop wireless network such as cellular networks or infrastructure based wireless LANS. Thus, a lot of research has been carried out to develop new MAC protocols for WMNS. In parallel with these efforts, several standards groups, in particular, IEEE 802 standard committees are driving the standardization of the technologies for WMNS in all areas ranging from personal area networks, local area networks, and metropolitan area networks, to even larger scale networks. the MAC protocols for WMNS can be classified into two categories: single-channel and multichannel MAC protocols which we will cover in the next sections is zero, then transmission can be started from this node, since the destination node will still be able to receive a packet correctly.

Improve virtual carrier sense. Virtual carrier sense can effectively reduce hidden nodes, but also cause more exposed nodes. In order to reduce the number of exposed nodes, directional virtual carrier sense is needed. a directional virtual carrier sense is proposed in to ensure that the operation of virtual carrier sense based on request to send/clear to send (rts/cts) matches the scenarios when both directional and Omni-antennas exist in the same network. However, when all nodes use Omni directional antennas, directional virtual carrier sense schemes similar to directional backoff needed to be developed. Such schemes rely on the availability of topology information, and cooperation between neighboring nodes.

• Dynamic tuning of backoff procedure. The backoff procedure can be modified in different ways. First of all, a different backoff instead of binary exponential backoff can be used. However, it is not preferred, since it is not compatible with csma/ca specified in IEEE 802.11. Another scheme is to assign different minimum and maximum contention windows for different nodes in the network. However, how effective this scheme can be to improve throughput performance is questionable. a scheme that dynamically tunes the contention window is proposed . in this scheme, the backoff is approximated by p-persistent backoff. Based on this model and also the estimated number of active stations in the network, the optimal persistence factor pmin is determined. With pmin, the contention window is computed as 2/pmin − 1. Simulations showed that this scheme could effectively improve throughput performance of csma/ca. in this scheme, each node is assumed to have packets to send following the poison process. In addition, the active

stations can be estimated. Moreover, the optimal persistent factor can be calculated based on the estimated number of active stations, estimated collisions, estimated idle periods, and so on. all these assumptions do not really match a real network, in particular in a WMN environment. We establish the existence of, and analyze the extent of, starvation in two-hop and three hop scenarios where a node is at most two or three hops from the GW. Specifically, we consider the interference model where some adjacent nodes are within carrier sensing range but are not within transmission range of each other. This is a common situation when a WMN is employed to provide network connectivity in dense environments such as a university campus. The contributions we propose a simplified Markov chain model to analyze WMNS with linear topology and numerically compute the degree of unfairness between nodes.

III.   Fair Binary Exponential Backoff Algorithm:

We propose a fair binary exponential backoff (FBEB) algorithm to reduce the extent of starvation. Our proposed algorithm uses the intuition that, to improve fairness, Mesh nodes that have successfully transmitted a data packet should not be permitted to eagerly transmit more data packets. By delaying the transmission of successive packets we are able to reduce the degree of starvation. This effect is achieved by adapting the contention window

For transmissions in the IEEE 802.11 protocol.[12]
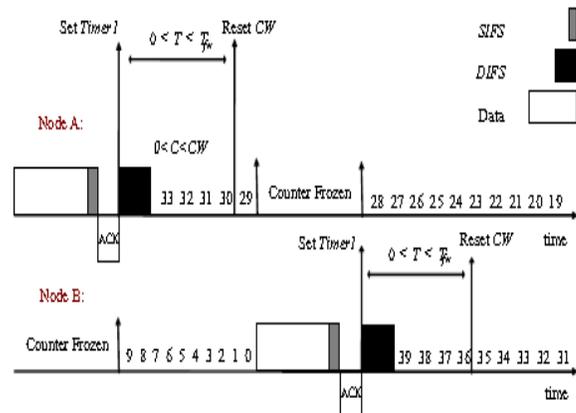
Algorithm 1 Fair Binary Exponential Backoff (FBEB) algorithm: Executed by each node $n \in \mathcal{N}$ in the network.

```
1:  T_fw = 5;
2:  K_succ = 4;
3:  DIFS = 2 timeslots + SIFS;
4:  CW_min = 31;
5:  CW_max = 1023;
6:  CW_n = CW_min;
7:  k_n = 0;
8:  while node n has packet to send do
9:      repeat
10:         Sense the channel continuously;
11:     until (The channel is idle for a DIFS)
12:     Choose C randomly in (0, CW_n);
13:     while (The channel is idle) do
14:         Decrease C;
15:     end while
16:     Send packet;
17:     flag = 1;
18:     while flag do
19:         if ACK received then
20:             Choose T randomly in (0, T_fw);
21:             Set Timer1 to T;
22:             k_n = k_n + 1;
23:             if k_n > K_succ then
24:                 Reset Timer1
25:             end if
26:             flag = 0;
27:         else if ACK timeout then
28:             k_n = 0;
29:             if CW_n < CW_max then
30:                 CW_n = 2 × CW_n;
31:             else
32:                 CW_n = CW_min;      //Reset CW_n
33:                 Drop the packet;
34:             end if
35:             flag = 0;
36:         end if
37:     end while
38: end while
Timer1 Reset Procedure:
1:  CW_n = CW_min;      //Reset CW_n
```

Fig: Fair Binary Exponential backoff Logarithm[12]

IV.   CONCLUSION:

Since many data transmission sessions in the Internet do not last very long, short-term starvation can dramatically degrade the quality of service provided by the network. We proposed a simple MAC layer technique, called FBEB algorithm, which was shown to alleviate the starvation in short-term sessions. Our proposed FBEB algorithm can increase the channel usage by the nodes farther to the GW about 7 times at the expense of losing 20% in total throughput. The proposed FBEB algorithm also resulted in a higher fairness index when compared with the CSP and Idle Sense schemes.

V.   Acknowledgements:

VI.   REFERENCES:

[1] M. Garetto, T. Salonidis, and E. W. Knightly, "Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks," in *Proc. of IEEE Infocom*, Barcelona, Spain, Apr. 2006.

[2] A. Margolis, R. Vijayakumar, and S. Roy, "Modelling throughput and starvation in 802.11 wireless networks with multiple flows," in *Proc. Of IEEE Globecom*, Washington, DC, Nov. 2007.

[3] C. Hua and R. Zheng, "Starvation modeling and identification in dense 802.11 wireless community networks," in *Proc. of IEEE Infocom*, Phoenix, AZ, Apr. 2008.

[4] G. Anastasi, E. Borgia, M. Conti, and E. Gregori, "IEEE 802.11 ad hoc networks: Performance measurements," in *Proc. of Int'l Conf. on Distributed Computing Systems*, Providence, RI, May 2003.

[5] M. Garetto, J. Shi, and E. W. Knightly, "Modeling media access in embedded two-flow topologies of multi-hop wireless networks," in *Proc. of ACM MobiCom*, Cologne, Germany, Aug./Sept. 2005.

[6] H. Y. Hsieh and R. Sivakumar, "IEEE 802.11 over multi-hop wireless networks: Problems and new perspectives," in *Proc. of IEEE VTC*, Birmingham, Al, May 2002.

[7] K. Xu, M. Gerla, L. Qi, and Y. Shu, "Enhancing TCP fairness in ad hoc wireless networks using neighborhood RED," in *Proc. of ACM MobiCom*, San Diego, CA, Sept. 2003.

[8] V. Gambiroza, B. Sadeghi, and E.W. Knightly, "End-to-end performance and fairness in multihop wireless backhaul networks," in *Proc. of ACM MobiCom*, Philadelphia, PA, Sept. 2004.

[9] J. Shi, O. Gurewitz, V. Mancuso, J. Camp, and E. W. Knightly, "Measurement and modeling of the origins

of starvation in congestion controlled mesh networks," in *Proc. of IEEE Infocom*, Phoenix, AZ, Apr. 2008.

[10] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: An optimal access method for high throughput and fairness in rate diverse wireless LANs," in *Proc. of ACM SIGCOMM*, Philadelphia, PA, Aug.2005.

[11] A. Raniwala, D. Pradipta, and S. Sharma, "End-to-end flow fairness over
IEEE 802.11-based wireless mesh networks," in *Proc. of IEEE Infocom*, Anchorage, AK, May 2007.

[12] Flow Starvation Mitigation for Wireless Mesh Networks Keivan Ronasi, Sathish Gopalakrishnan, and Vincent W.S. Wong Department of Electrical and Computer Engineering The University of British Columbia, Vancouver, Canada e-mail: {keivanr, sathish, vincentw}@ece.ubc.ca

Sk.Mansoor Rahaman received b.tech degree from affiliated college of JNTU Hyderabad in 2010 and he is currently pursuing his master's degree in computer networks and security from K.L.University (2010-2012). His research interest are wireless mesh networks and wireless sensor networks.

Venkaiahnaidu.Adapa received b.tech degree from JNTU kakinada affiliated college of JNTU Kakinada in 2009 and he is currently pursuing his masters degree in computer networks and security from K.L.University (2010-2012). His research interest are cognitive wireless mesh networks.

Cluster Formation in Manet Using S/N Ratio of Channel Connecting Two Nodes.

DEEPA NAIK

The configuration of mobile ad hoc networks (MANETs) is constantly changing. The attacks by malicious nodes or any faults in network hardware or software directly affects the network resources specially Bandwidth. This affects network performance. It is proposed a cluster scheme purely based on bands of S/N ratio .

Key Words
S/N ratio, migration, Cluster Head

The concept:
Here an area to be serviced divided into regular shaped cells. Each of these cells is assigned a range of S/N ratio band. The same S/N ratio band cannot be used in other cells. Here each node align itself different clusters. As the nodes move around, they would change from cluster to cluster .The nodes can find out which cluster is governed by what S/N ratio band... When they did not receive a signal from any of the nodes in a cluster, they would try other clusters until they found one.
Effect of faulty nodes is waste of Bandwidth..

1. Introduction
Bandwidth is a limited resource in MANETs. To prevent unnecessary allocation of Bandwidth for non performing nodes we have deviced an algorithm for the prevention of wastage of Network bandwidth **.**
The cluster formation: Every group of node is formed together and the arranged in one Group. The main purpose of cluster formation is the reduce the transfer Rate which will reduce propagation delay and bandwidth consumption. For the formation of clusters various criteria have been developed.

Here I am using Cluster making algorithm based on S/N ratio
per cluster we have to limit the number of members.

S/N ratio= Power of Signal in the channel connecting two nodes/Power of noise in the channel connecting two nodes.

S/N in dB = 20 Log S/N ratio.

We will be using Cluster making algorithm based on S/N ratio .Here we assume a Mannet as a network distributed over land areas called Clusters , each hypothetically controlled by by at least one fixed-location Cluster head (base station).When joined together these clusters provide radio coverage over a wide geographic area. This enables a large number of nodes to communicate with each other and with fixed cluster heads anywhere in the network, via cluster heads, even if some of the nodes are moving through more than one cluster l during operation...

This will offer a number of advantages over alternative solutions:

- increased capacity
- reduced power use
- larger coverage area
- reduced interference from other signals

The concept:

Here an area to be serviced divided into regular shaped cells. Each of these cells is assigned a range of S/N ratio band. The same S/N ratio band cannot be used in other cells. Here each node align itself different clusters. As the nodes move around, they would change from cluster to cluster .The nodes can find out which cluster is governed by what S/N ratio band... When they did not receive a signal from any of the nodes in a cluster, they would try other clusters until they found one.

Effect of faulty nodes is waste of Bandwidth..

Principle: The communication between any nodes in Manets is very much noise affected and hence the S/N ration comes in to play. Hence there will be different S/N ratios for different pair of nodes. We propose a clustering scheme between nodes having S/N ratio in a band and the nodes having heighest S/N will be selected as Cluster heads1 and 2 .

2. In our system nodes organize in clusters only difference between clusters is S/N ratio band. Clusters we propose here are very similar to Cells in mobile telephony. Nodes with failures will have extremely low S/N ratios with any o

These nodes. (Literally zero).The malicious nodes will constantly prey on victim nodes and constantly changing hence will have constantly changing S/N ratios. This will be detected with algorithm... **3**

Evolution of clusters:

When two nodes in Manets communicate they will search simultaneously search for neighboring nodes. They notify if any such nods and the S/N ratio of their respective channels.

N3   **Nfaulty**

N1
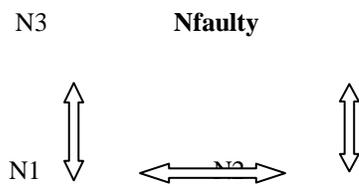
From Node	To Node	S/N ratio	Compare if they fall in a narrow band, €
N1	N2	Db(n1n2)	
N1	N3	Db(n1n3)	Is Db(n1n2)- Db(n1n3)< €1
N2	N4	Db(n2n4)	Db(n2n4)- Db(n1n3)< €1
N2	N5	Db(n2n5)	Db(n2n5)- Db(n1n3)< €1
N2	Nfaulty	Db(n2nf)	Db(n2nf)> €1
N2	Nx(a genuine node but constantly changing its location and hence its Db)	Dbx=Will wait for a suitable cluster with its Db fits in.	The performance of this node may not be effective or useful to the overall network performance.

For cluster 2 another band €2 is defined.Simillarly for other clusters different bands can be selected. Say €1=10 dB,€2=15dB,€3=20dB.

N1,N2,N3,N4…..NT be the nodes,

Let Db(nanb) be the S/N ratio between two pairs of nodes a and b.

Hence Db (n1n2), Db (n1n3), Db (n2n4), between nodes N1, N2 etc.

.Let Db (n2nf) be S/N ratio from N2 to faulty, Let Db (n2nNe) be S/N ratio from N2 to Non efficient malicious node. Let Dab be S/N ratio of a constantly changing node. The node having dynamically varying S/N ratio will ultimately settle in one or the other cluster. i.e. Dbx=Will wait for a suitable cluster with its Db fits in.

1. Calculate S/N for each pair of nodes.

2. Assign pair of nodes having S/N in a band €1 in a

Cluster K1.Assign other nodes with same band to same cluster till total number of nodes in a cluster is less than or equal to 'L'.Similarly assign pair of nodes having S/N in a band €2 in a cluster K2.

3.Suppose a node Ng fits in both clusters say K1 and K2 but with different other nodes say Nh and Ni. Then the first calculation will be taken in to account.

4. What if a node falls into band €1 and the clusterK1 is already populated. (ie more than 'L' maximum limit).It will form a standalone Cluster. And whenever there is a vacancy it will fit in. As it will be assigned seniority malicious nodes cannot enter the vacancy easily.

5. What if a node does not fall into any clusters. Then also it will form a standalone Cluster. And whenever there is a S/N match it will fit in. As it will be assigned seniority malicious nodes cannot enter the vacancy easily.

6. The first two nodes will be assigned cluster head1 and cluster head2.They will assign management of network resources. Hence malicious node has to enter and form a cluster instantly in order to head such a cluster which is a remote possibility. Because it has to first establish a communication with network and by that time the elapsed time for cluster making will fulfill.

Previous work: Cluster head selection based on heighest degree:It is based on the number of neighborhood nodes a particular node is having around.

Whenevr the election procedure is needed the node broadcast their identifier which is assumed to be unique in the same network. According to the number received identifiedr (ID) of every node compuetes it degree .One having maximum degree becomes the cluster head.This algorithm fails because ther degree of the node changes very frequently. The cluster heada are unlikely to play their role for a long time . This may affect stability in cluster management .

The Lowest load principle also known as Identifier based node selection.It is better than the heighest degree in terms of throughput. Major drawback of this algorithm is the smallest ID's which may lead to batter drainage of certain nodes.

Migration from one cluster to another and node alignment
When a node is moved away from a first cluster and closer to a second cluster, the node listens to channels and switches itself to another cluster. The new cluster automatically selects the node which will align to it only if there is a vacancy. The migrated node automatically switches from the current cluster to the new cluster and communication continues.If theire is no vacancy it will form a standalone cluster.If there is a message not completely sent by migrating node it simultaneously moves to new cluster at the same time and sends the remaining part of the message with the strongest signal.

Effect of migration on cluster stability and management:

The effect of migration of nodes is having two effects on cluster.

1. On the stability of clusters: The nature of AdHoc network it itself is constantly changing. Hence any scheme is bound to have this limitation.Here migration is not dependent on physical distance but on the stengngth ogf the channel between migrating node and member nodes in a cluster.
2. On the Cluster head: The mannets are continuously evolving. Hence if a cluster head1 itself migrates the cluster head 2 will take over.

   Here it is inherent that the cluster Heads should have more network resources and battery power. If Cluster head 1 finds that it is running out of battery or memory than it recluses itself from heading the cluster and braodcast a message of Cluster head2 taking over.
3.
4. Meanwhile Cluster head broadcasts a message about the Cluster head 2 taking over.
5. Range of dBs:

Say: range of dBs

i.   5-15 db
ii.   20-30 dB
iii.   35-45 dB
iv.   40-50 dB
v.   55-65 dB

1. First any two nodes will establish communication and their channel S/n may be anywhere in above ranges.

A   B
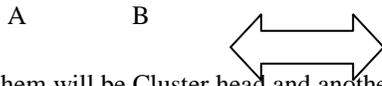
One of them will be Cluster head and another is sub cluster head for the cluster formulating subsequently.

2. Maximum number of nodes in a cluster is to be limited for the sake of cluster management. Say maximum 5 nodes in a cluster.

So if more more than 5 nodes fall in a band of cluster first five will form a cluster. The sixth will be a stand alone cluster connected to the earlier cluster.
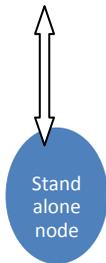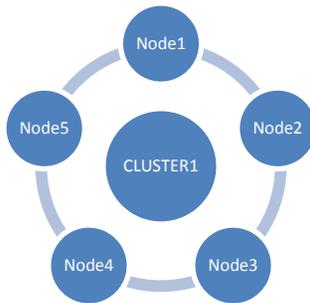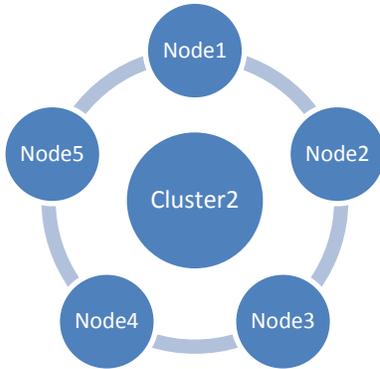
3. What if a node is having an in-between dB range say 17 dB.

   It is connected to some member of a cluster already formed and it will form a standalone node.

4. Standalone nodes although increase cluster overhead but they may form a connection in between clusters of two different dB range.
5. Network overhead wrt number of clusters.

   Suppose n is even saying 10. The number of clusters formed is limited by closeness of db ranges.

   Nodes are A,B,C,D,E,F,G,H,I,J.

   Let us study propagation delay in sending a message from A to J.

   All are standalone nodes clusters, hence maximum 5 clusters

   AB-----CD-----EF-------GH-----IJ 4 Propagation delay

3 Nodes in a cluster.

   ABC-------DEF------GHI-------J   3 pd

   ABC------DEF-----GH----IJ   3 pd

   ABC---------DE------FG----------IJ

4 nodes in a cluster

   ABCD---------EFGH----------IJ 2 pd

   ABCD-----EFG-----HIJ 2 pd

   ABCD------EFG-----H----IJ

   ABCD------EF-----GH-----IJ---

5 nodes in a cluster:

ABCDE------FGHIJ

Less propagation delay is possible when only number of nodes in a cluster is n/2. But we can not allow more than a certain number of nodes in a cluster.Hence cluster management overhead and propagation delay(parameter of QOS) to be suitably optimized. Number of clusters will be slowly formed when' n' is small. Cluster making is difficult. E.g. Ships of a naval force in an ocean.But with mannet having more nodes cluster will be evolved fast. There is a limitation in that probability of migration of nodes from one cluster to another is also high which will have an effect on cluster stability.

6. To increase better utilization of cluster resources we have to increase the number of nodes forming the cluster. For that we have to increase the dB band. Say we have

A -----B = 200 dB : A -----C = 20 dB

Now to include all these three we have to make a band of 10 -250 dB.

Suppose B is sending a message to C via A. In order to form a cluster it has to compromise with QOS from A to C.

We assume that anyway the communication have to be carried out and there is no alternate route from B to C we have to compromise this QOS.

5. Conclusion
Our current research work on cluster formation is in accordance with dynamic nature of mannets.It is prone to suffer from fault Occurrences in harsh environments because of constantly changing signal power.

6. Future Work
Following this conceptual framework we expect to Carry out simulations of the proposed clustering scheme Using network simulator NS-2. We would like to optimize network resources as we have assigned Cluster head 1 and Cluster head2 .
.

Heuristic Approach for Secure Energy Efficient Routing in Mobile ad hoc Network

Nithya.S[1] and Chandrasekar.P[2]

[1]II ME (CS), Sri Shakthi Institute of Engineering and Technology, Anna University of Technology, Coimbatore
[2] Asst Professor (S) ECE, Sri Shakthi Institute of Engineering and Technology, Anna University of Technology, Coimbatore

ABSTRACT

In today's scenario, energy savings and security is the major problem in all kind of the networks. MANET is a network, which is very popular due to its unique characteristics from all the other types of networks. MANET is a network having tiny light weighted nodes, with no clock synchronization mechanisms. Generally, in this type of network the exhaustion of energy will be more and as well, the security is missing due to its infrastructure less nature. Due to the lack of energy, the link failure may occur and the network lifetime also gets affected. Similarly, the node causes cheating during the transmission process in the network, most MANET routing protocols are vulnerable to attacks that can freeze the whole network. Thus these may affects the performance of the network. To overcome these problems, we propose a new secured energy efficient routing algorithm. This algorithm holds two mechanisms. Initially it makes all the active state nodes to sleep when not in use, and then finds the efficient path for reliable data transmission. Secondly, provides the security against attacks using a new cryptographic mechanism. By simulation based studies, we show that this algorithm effectively provides higher security, less energy utilization, less overhead and less end to end delay.

Keywords— Energy, Link failure, MANET, Network, Security.

I. INTRODUCTION

Wireless ad hoc networks draw lots of attentions in recent years due to its potential applications in various areas. Among the various network architectures, design of the mobile ad hoc networks (MANET) plays an important role. Such a network can either operate in a standalone fashion with the ability of self-configuration and no clock synchronization mechanism. Mobile Ad-hoc networks are self-organizing and self-configuring multi-hop wireless networks where, the structure of the network changes dynamically. No base stations are supported in such an environment, and mobile hosts may have to communicate with each other in a multi-hop fashion. Minimal configuration and fast deployment make MANETs suitable for emergency situations like natural or human-induced disasters and military conflicts. The performance of a mobile ad hoc network mainly depends on

the routing scheme. Our critical issue for almost all kinds of portable devices supported by battery power is power saving. Without power, any mobile device will become useless. Battery power is a limited resource, and it is expected that battery technology is not likely to progress. Hence lengthen the lifetime of the batteries is an important issue, especially for MANET, which is all supported by batteries [1],[2],[3],[4]. The previous energy-efficient algorithms can try to reduce the energy consumption. However while considering minimum energy path, they do not considering the reliability of the links. This may result in low quality of service, less reliable path. When we consider the reliability of the network, energy consumption of the network will be high. Similarly, Security is a more sensitive issue in MANETs than any other networks due to lack of infrastructure and the broadcast nature of the network. The nature of ad hoc networks poses a great challenge to system security designers due to the following reasons: Firstly, wireless network is more susceptible to attacks ranging from passive eavesdropping to active interfering, Trusted Third Party adds the difficulty to deploy security mechanisms, mobile devices tend to have limited power consumption and computation capabilities, finally, node mobility enforces frequent networking reconfiguration which creates more chances for attacks. There are five main security services for MANETs: authentication, confidentiality, integrity, non-repudiation, availability. Among all the security services, authentication is probably the most complex and important issue in MANETs. Several security protocols have been proposed for MANETs, there is no approach fitting all networks, because the nodes can vary between any devices. In order to overcome these problems, we propose a new secured energy efficient routing algorithm. The main contribution of this paper is in showing how power aware routing must not only be based on node specific parameters (e.g. residual battery energy of the node), but must also consider the link specific parameters (e.g. channel characteristics of the link) as well, to increase the operational lifetime of the network. And also provides the security against route reply attacks using a check sum mechanism. It may also balance the traffic load in the network, while finding the reliable transmission path. Sleep/Active mode approach and Transmission Power Control Schemes are the main two methodologies, which are mainly responsible for considerable energy saving.

II. AN OVERVIEW OF RELATED WORK

Mobile devices coupled with wireless network interfaces will become an essential part of future computing environment consisting of infra-structured and infrastructure-less mobile networks. Wireless local area network based on IEEE 802.11 technology is the most prevalent infra-structured mobile network, where a mobile node communicates with a fixed base station, and thus a wireless link is limited to one hop between the node and the base station. Mobile ad hoc network (MANET) is an infrastructure-less multi hop network where each node communicates with other nodes directly or indirectly through intermediate nodes. Thus, all nodes in a MANET basically function as mobile routers participating in some routing protocol required for deciding and maintaining the routes.

Since MANETs are infrastructure-less, self-organizing, rapidly deployable wireless networks, they are highly suitable for applications involving special outdoor events, communications in regions with no wireless infrastructure, emergencies and natural disasters, and military operations. Routing is one of the key issues in MANETs due to their highly dynamic and distributed nature. In particular, energy efficient routing may be the most important design criteria for MANETs since mobile nodes will be powered by batteries with limited capacity. Power failure of a mobile node not only affect the node itself but also its ability to forward packets on behalf of others and thus the overall network lifetime.

Energy management in wireless networks is very important due to the limited energy availability in the wireless devices. It is important to minimize the energy costs for communication as much as possible by practicing energy aware routing strategies. Based on the observations of signal attenuations, many routing protocols are operated. Energy aware routing algorithm would select a route comprising multiple short distance hops over another one with a smaller hop count but larger hop distances.

The PAMAS (Power aware Multi access protocol with signaling) [5] protocol allows a host to power its radio off when it has no packet to transmit/receive or any of its neighbors is receiving packets, but a separate signaling channel to query neighboring hosts' states is needed.

In PAMAS, [5] they provide several sleep patterns and it allows the mobile nodes to select their sleep patterns based on their battery power. But this needs a special hardware called RAS (Remote Activated Switch. But they biased towards smaller hops typically led to the selection of paths with a very large hop count.

The PARO [6], [7] has proposed for the situation where the networks having the variable transmission energy. This protocol essentially allows an intermediate node to insert itself in the routing path if it detects potential savings in the transmission energy.

Later, Connected-dominated set based power saving protocol is proposed. In which some hosts must as a coordinators, which are chosen according to their remaining battery energies and the numbers of neighbors they can connect. In this type of network, only coordinators need to awake, other hosts can enter the sleeping mode.

Min-Hop routing is the conventional "energy unaware" routing algorithm, where each link is assigned based on the identical cost. In which it simply selects the routes based upon the number of hops. Less number of hop counts path is considered as a route for transmission of packets. Thus results in less reliability and power wastage. Min Energy routing is another power aware routing algorithm, which simply selects the path corresponding to the minimum packet transmission energy for reliable communication, without considering the battery power of individual nodes. In which the number of hops and delay increases. This results in less energy consumption but with less reliability

MMBCR [4] is a power-aware routing algorithm, which selects the path whose critical node has the highest residual battery energy. The node having the battery level, greater than the threshold value is considered as the strongest node. It selects the route with strongest node for packet transmission. Since the MMBCR algorithm never tries to minimize the total transmission energy along a path, it can lead to overall higher energy consumption and consequently, a reduction of the average node lifetime.

CMMBCR[4] algorithm uses the minimum energy path initially, as long as the battery power level on all the nodes in the selected path lies above a certain threshold. Once one or more of nodes on all possible paths falls below this battery protection threshold, the algorithm switches to the MMBCR mode. This results in overall higher energy consumption but provides higher reliability. MRPC (Maximum Residual Packet Capacity) and CMRPC (Conditional MRPC) were proposed. In this power aware routing algorithm it considers residual battery energy of the nodes along with link reliability for route selection. It uses Min-max formulations which is similar to MBCR and MMBCR.

Huaizhi Li and Mukesh Singhal [8] have presented an on-demand secure routing protocol for ad hoc networks based on a distributed authentication mechanism. The protocol has made use of recommendation and trust evaluation to establish a trust relationship between network entities and it uses feedback to adjust it. The protocol does not need the support of a trusted third party and it discovers multiple routes between two nodes. Sec AODV [9] is the one of the protocol that incorporates security features of non-repudiation and authentication, without relying on the availability of a Certificate Authority (CA) or a Key Distribution Center (KDC). They have presented the design and implementation details of their system, the practical considerations involved, and how these mechanisms are used to detect and thwart malicious attacks.

Packet conservation Monitoring Algorithm (PCMA) [10] can be used to detect selfish nodes in MANETs. Though the protocol addresses the issue of packet forwarding attacks, it does not address other threats

Syed Rehan Afzal et al. [11] have explore the security problems and attacks in existing routing protocols and then

they have presented the design and analysis of a secure on-demand routing protocol, called RSRP which has confiscated the problems mentioned in the existing protocols. Moreover, unlike Ariadne, RSRP has used a very efficient broadcast authentication mechanism which does not require any clock synchronization and facilitates instant authentication.

A credit-based Secure Incentive Protocol (SIP) is used to stimulate cooperation in packet forwarding for infrastructure less MANETs. Though the protocol addresses the issue of packet forwarding attacks, it does not address other threats.

III. SECURE ENERGY AWARE ALGORITHM

MER is an algorithm which provides less energy consumption but results in less reliability. As in case of MMBCR [4] algorithm, results in high reliability but with higher energy consumption. A mobile node consumes its battery energy not only when it actively sends or receives packets but also when it stays idle listening to the wireless medium for any possible communication requests from other nodes. Thus, energy efficient routing protocols minimize either the active communication energy required to transmit and receive data packets or the energy during inactive periods. Our newly proposed secure energy aware algorithm holds two mechanisms.

Mechanism 1: This mechanism deals with the reduction of energy consumption. It makes all the active state nodes to sleep when not in use by means of active sleep state methodology. This Active /sleep state methodology initially categorize the energy as active communication energy and inactive communication energy. The active communication energy was reduced by adjusting the power of the each node to reach only the particular destination and not more than that. The inactive communication energy was reduced by simply turns off the node during the idle case.

This leads to considerable energy savings, especially when the network environment is characterized with low duty cycle of communication activities. Secondly, it will find the route with least cost path based on the reliability and the residual battery energy. This algorithm assumes RREQ (Repeat Request) for reliable packet transmission in each hop. If the packet or its acknowledgement is lost, the sender will retransmit the packet. To formulate this algorithm, assume E be the energy expected by the node to transmit the packets from source to destination.

E-> Expected Energy to Transmit a Packet
B-> Total Residual Battery Energy
R= B-E-> Remaining Residual Battery Energy

The ratio of the fraction of residual battery energy to be consumed to the total residual battery energy (B) gives the link weight. The path with less weight is to be selected. The Link weight is defined as the fraction of the residual battery

energy that node i consumes to transmit a packet reliably over (i, j). Link weight is determined using Dijkstra's algorithm.

If the residual energy of the nodes is not considered, then the energy in the best path's node will be consumed more than the other nodes in the network. In this model the consumed energy by a node during packet transmission consists of two elements. The first element is the energy consumed by the processing part of the transceiver circuit, and the second element is the energy consumed by the transmitter amplifier to generate the required power for signal transmission.

In Ad-hoc network, the packets are transmitted with minimum power, which is required for decoding the packets. In such a situation, TPC (Transmission Power Control) scheme is used. This transmission power control approach can be extended to determine the optimal routing path that minimizes the total transmission energy required to deliver data packets to the destination.

In wireless communication transmission power has strong impact on bit error rate, and the inter radio interference. Thus this transmission power control scheme which will adjust the transmission power of the node based on the link distance. If TPC is not present, then the maximum transmission power is utilized.

Mechanism 2: This mechanism deals with the security aspects. In order to make our proposed algorithm more secure, a new cryptographic check sum mechanism is used. The proposed algorithm is very effective as it detects the malicious node quickly and it provides security against the attacks. Among all the security services, authentication is probably the most complex and important issue in MANETs. Cryptographic mechanisms make use of a hash code. Hash code does not use a key but is a function only of the input message. The message plus concatenated hash code is encrypted using symmetric encryption.
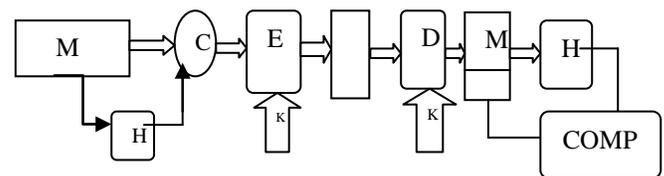

Fig 1: Basic Security Function

M-Message
H-Hash Code
C-Concatenation
E-Encryption
D-Decryption
K-Secret Key
COMP-Comparison

The figure indicates that, the message and the hash function are concatenated. Then the concatenated hash code along with the message is encrypted using symmetric key

encryption. The bank block indicates the encrypted value of concatenated hash code with message. The Message must be transferred only between the source and the destination using the secret key, thus the data transmission is more secure and has not been altered.

The comparative block predicts the absolute key value with secured message. The hash code provides the structure or redundancy required to achieve authentication. Because encryption is applied to the entire message plus hash code, Confidentiality is also provided.

If in the case of many intermediate states present between the source and the destination, then the security is achieved by means of digital signatures.

Thus, our new secure energy aware algorithm with these two mechanisms enhances the routing problem and manages the network resources of achieving fair resources usage across the network node with higher security.

IV. PERFORMANCE EVALUATION:

1. Simulation model:
Consider an ad hoc network in which nodes are uniformly distributed in a square area. In the network, sessions are generated between randomly chosen source-destination nodes with exponentially distributed inter-arrival time. The source node of the session transmits data packets with the constant rate 1 packet/sec. We developed our simulation model using ns 2.34 simulator. NS-2 simulator allows extracting from a simulation many interesting parameters, like throughput, data packet delivery ratio, end-to-end delay and overhead.

To have detailed energy-related information over a simulation, we modified the ns-2 code to obtain the amount of energy consumed over time by type (energy spent in transmitting, receiving, overhearing or in idle state). This way, we obtained accurate information about energy at every simulation time. We used these data to evaluate the protocols from the energetic point of view: we will see the impact of each protocol on different new parameters, like the number of nodes alive over time (to check the lifetime of nodes), the expiration time of connections (to see the network lifetime), and the energy usage divided by type (receiving, transmitting, overhearing).

1.1. Practical Considerations:
The routing protocols for MANET'S are generally categorized as table driven, and on demand driven based on the timing of when the routes are updated. RWMECR algorithm can be implemented with the existing routing protocols for ad hoc networks. Here, we implemented with AODV as the routing protocol. The algorithm performance was compared with the normal AODV protocol. An AODV is an on demand routing protocol that combines the capabilities of both DSR and DSDV protocol. It uses route discovery and route maintenance from DSR and in addition to the hop by hop routing sequence numbers and periodic beacons from Destination-Sequenced Distance vector (DSDV) routing protocol.

AODV is an on demand routing protocol in which routes are discovered only when a source node desires them. Route discovery and route maintenance are two main procedures: The route discovery process involves sending route-request packets from a source to its neighbor nodes, which then forward the request to their neighbors, and so on. Once the route-request reaches the destination node, it responds by uni casting a route-reply packet back to the source node via the neighbor from which it first received the route request.

When the route-request reaches an intermediate node that has a sufficiently up-to-date route, it stops forwarding and sends a route-reply message back to the source. Once the route is established, some form of route maintenance process maintains it in each node's internal data structure called a route-cache until the destination becomes inaccessible along the route. Note that each node learns the routing paths as time passes not only as a source or an intermediate node but also as an overhearing neighbor node.

Our simulation settings and parameters are summarized in table 1

Area Size	1000 X 1000
Simulation time	100 s
Number of Nodes	50
MAC type	MAC 802.11
Traffic Source	CBR
Initial Energy	1000 J
Packet Size	512
Routing Protocol	AODV

Table: 1

1.2 Simulation Results
The following results show the operation of new secure energy aware algorithm. Some parameters like packets received, Energy consumption per packet transmission, end to end latency and packet delivery ratio are analyzed to verify the performance of the new power aware mechanisms. As dealing with the energy aspect, AODV protocol shows good energy efficiency when compared with the all other existing protocols. Here, we compare our algorithm performance with the AODV protocol.

Energy Consumption per packet:
It defines the energy consumed by a node to transmit a packet from source to destination. In the below graph we compared the plain AODV protocol with our new secure energy aware mechanism. By means of new secure energy aware mechanism the power consumed by the node to transmit to the packet was decreased at a higher rate. The energy consumption per packet was decreased as previous. This will highly increases the network life time.
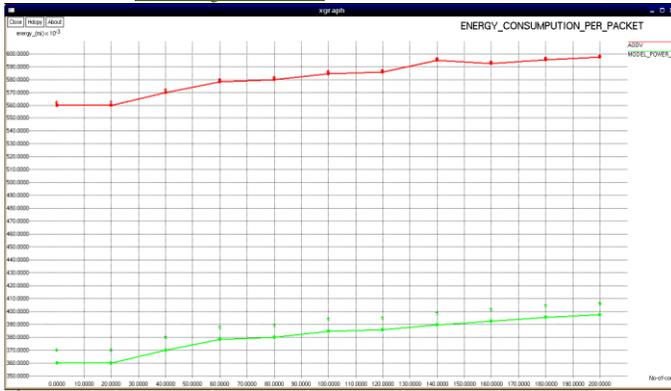
Fig 2: Energy Consumption per Packet

Figure 4: End To End Latency

Packet delivery ratio:

Data packet delivery ratio can be calculated as the ratio between the number of data packets that are sent by the source and the number of data packets that are received by the sink. This is the amount of successful received bits at the destination nodes for the entire simulation period. Packet delivery ratio should be always high for the efficient algorithm or a protocol. The below figure shows the packet delivery ratio was high when compared with the previous methodology.

The end to end latency of the new secure energy aware mechanism was highly reduced when compared with normal protocol operations.

Packets Received:

It denotes the amount of packets received by the destination during the simulation. By means of new algorithm, the amount of packets discarded or dropped was reduced. By means of our secure energy aware mechanism, we may increase the amount of packet received, below figure which shows that, the amount of packets received by the destination node with higher security.

Figure 3 Packet delivery ratio
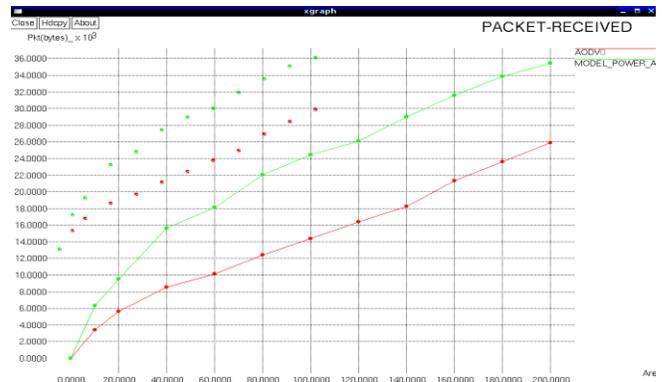
Figure 5: Total Packet Received

End To End Latency:

End-to-end Latency refers to the time taken for a packet to be   transmitted across a network from source to destination.End to end latency which includes all possible delays caused by buffering during route discovery time, queuing at the interface queue, retransmission, and processing time. It defines the ratio of interval between the first and the second packets to a total packets delivery. This below figure shows the result of end to end latency. For the good results, the End to End latency will be decreased

As dealing with the security aspects, we compare our new secure energy aware algorithm with Sec AODV and RSVP protocol.

Overhead:

The control overhead is defined as the total number of routing control packets normalized by the total number of received data packets. When compared with the other existing protocols, our mechanisms hold less number of overhead packets.

Figure 6: Overhead

Data Delivery:

Packet Delivery Ratio denotes the ratio of total number of packets received at the destination from the source. Whereas the Data Delivery denotes, the overall data information received by the destination with full security. When compared with secure AODV and RSVP, our secure aware algorithm (model AODV) provides higher data delivery.
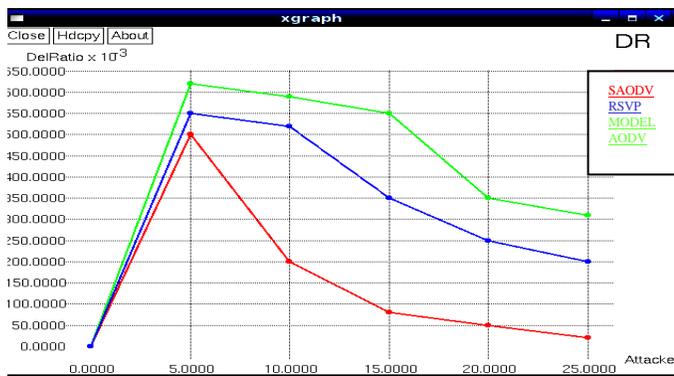

Figure 7: Data Delivery

V. CONCLUSION

In this paper, a new secure energy aware routing algorithm was proposed. It mainly defines the least cost path based on the reliability and the remaining energy of the node for packet transmission from, source to destination, and making the sleep/active state methodology for providing the energy efficiency. Later this algorithm provides a new cryptographic check sum mechanism to prevent the communications from attackers. By means of these features, we may effectively secure our data's with minimal energy consumption Thus; this algorithm can effectively reduce the energy consumed by the node as well as increases the security and reliability of the network. This in turn increases the operational lifetime and it maintains the load traffic as well.

REFERENCES

[1]   X.-Y. Li, Y. Wang, H. Chen, X. Chu, Y. Wu, and Y. Qi, "Reliable and energy-efficient routing for static wireless ad hoc networks with unreliable links," IEEE Trans. Parallel Distrib. Syst., vol. 20, no. 10, pp. 1408–1421, 2009.

[2]   B. Mohanoor, S. Radhakrishnan, and V. Sarangan, "Online energy aware routing in wireless networks," Ad Hoc Networks, vol. 7, no. 5, pp. 918–931, July 2009.

[3]   B.Chen , k.Jamieson, H.Balakrishnan, and r.Morris, "span:An energy Efficient cocrdination algorithm for topology Maintenance in ad hoc wireless networks", Proc of the international conference on mobile computing and networking , pp.85-96,2001.

[4]   C.k.Toh ,H.cobb and.Scott,"Performance Evaluation of battery life aware routing schemes for wireless adhoc networks", Proceedings of IEEE ICC'01 ,June 2001.

[5]   S. Singh and C.S. Raghavendra, "PAMAS-PowerAware Multi-Access Protocol with Signaling for Ad Hoc Networks", ACMCommunication Review, July 1998.

[6]   J. Gomez, A. T. Campbell, M. Naghshineh, and C. Bisdikian, "Paro: supporting dynamic power controlled routing in wireless ad hoc networks," Wireless Networks, vol. 9, no. 5, pp. 443–460, 2003.

[7]   J. Gomez, A. Campbell, M. Naghshineh and C. Bisdikian, "PARO: A Power-Aware Routing Optimization Scheme for Mobile Ad hoc Networks", draft-gomez-paro manet-00.txt, IETF, Work in Progress, February, 2001

[8]   Huaizhi Li and Mukesh Singhal, 2006."A Secure Routing Protocol for Wireless Ad Hoc Networks", in proceedings of 39th Annual Hawaii International Conference on System Sciences, Vol.9.

[9]   A. Patwardhan, J. Parker, M. Iorga, A. Joshi, T. Karygiannis and Y. Yesha, 2008. "Thresholdbased intrusion detection in ad hoc networks and secure AODV", Vol.6, No.4, pp.578-599.

[10] Tarag Fahad & Robert Askwith, 2006. "A NodebMisbehaviour Detection Mechanism forbMobile Ad-hoc Networks" The 7th Annual PostGraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting

[11] Syed Rehan Afzal, Subir Biswas, Jong-bin Koh, Taqi Raza, Gunhee Lee, and Dong-kyoo Kim, 2008. "RSRP: A Robust Secure Routing Protocol for Mobile Ad hoc Networks", IEEE Conference on Wireless Communications and Networking, pp.2313-2318.

Estimation of optimal cutting parameters of plane turning using quantum inspired evolutionary algorithm

R.S.S. Prasanth[1], K. Hans Raj[2]

*(Technical College, Dayalbagh Educational Institute, India)
** (Faculty of Engineering, Dayalbagh Educational Institute, India)

ABSTRACT

Turning is a versatile machining process that involves different cutting parameters and conditions. The surface finish is the vital design requirement as it is a key indicator of quality of the work piece. This work, presents the application of Quantum Inspired Evolutionary Algorithm (QIEA), that essentially exploits some principles of quantum mechanics such as Q-bits, superposition, quantum gate and quantum measurement, for the process optimization of plane turning. The QIEA estimated optimal cutting parameters i.e., cutting speed, feed rate, tool nose radius and depth of cut of plane turning for improved surface finish within the operating conditions. The results are compared with real coded genetic algorithm (RCGA) and differential evolution algorithm (DEA). The results obtained by Quantum Inspired Evolutionary Algorithm are better than those reported with RCGA and are comparable to those of DEA.

Keywords – **Differential evolution, Plane turning, Quantum Inspired Evolutionary Algorithm, Surface finish**

I. INTRODUCTION

The key attribute of any machined part is its surface finish, which is a technical requirement as the machined part necessarily have to interact with other parts of the larger mechanical system per se. Surface finish is actually the degree of smoothness of a machined part, which is the result of the surface roughness. Surface roughness is undesirable, but difficult and expensive to control during manufacturing. Decreasing roughness of a surface will usually exponentially increase its manufacturing costs. Optimizing any machining process in want of better surface finish using experimental methods is very difficult and cost intensive, and often fail to achieve good repeatable optimal or near optimal results. Plane turning is an indispensible metal removing process that finds wide applications in all the manufacturing industries. In plane turning, the cutting tool establishes contacts with workpiece, at a single point, and resulting, heat and wear at the contact point between cutting tool and workpiece. Consequentially, tool life gets affected and surface roughness increases. Thus, it is imperative to select optimal cutting parameters, such as cutting speed, feed rate, and depth of cut, that are known to have a significant impact on surface quality of the workpiece [1].

To address the limitations of laborious, cost intensive traditional techniques, soft-computing techniques are increasingly inviting the attention of researchers, as they are capable of handling highly non linear complex real world machining optimization problems [2]. The problem of plane turning process optimization was attempted using binary coded genetic algorithms (BCGA) to estimate optimal cutting conditions for the process [3, 4]. An empirical surface roughness model of plane turning was enumerated and solved using real coded genetic algorithm (RCGA), which does not suffer from imprecision and premature convergence, unlike BCGA [5, 6].The differential evolution algorithm (DEA) was also applied on the same optimization problem in order to reduce the surface roughness [7]. Although Quantum Inspired Evolutionary Algorithm was firstly introduced by Narayanan and Moore to solve TSP [8], in which the crossover operation was performed based on the concept of interference. Ever since, Han and Kim exploited the quantum mechanics principles such as Q-bits, superposition, quantum gates and quantum probabilistic measurement and developed a more practical

algorithm, Quantum Inspired Evolutionary Algorithms gained greater attention of the scientific fraternity.

Quantum Inspired Evolutionary Algorithm was applied on some engineering optimization problems. However, the problem of process optimization of turning was not solved yet, using Quantum Inspired Evolutionary Algorithm. On the other hand, Gexiang Zhang, reported that, although Quantum Inspired Evolutionary Algorithm is reportedly better than genetic algorithm (GA), there are a few comparisons made between Quantum Inspired Evolutionary Algorithm and Differential Evolution Algorithm [11]. Therefore this work presents the performance comparison of Quantum Inspired Evolutionary Algorithm on the process optimization of plane turning. The paper is further organized as follows. In section II, the empirical model of plane turning is presented. Quantum Inspired Evolutionary Algorithm is explained in detail in section III. In section IV results of the experiments conducted on Quantum Inspired Evolutionary Algorithm for process optimization plane turning are compared with the results of Differential Evolution Algorithm. The conclusions are presented in section V.

II. MODEL OF PLANE TURNING

The problem of prediction of optimal cutting parameters for plane turning may be enumerated as objective minimization problem as

$$\text{Min} \quad R_a \ (v, \ f, \ d, \ r)$$

The average surface roughness Ra [6] is calculated by the following empirical formula

$$\mathbf{R_a = (1.0632 \ x \ f^{1.0198} \ x \ d^{0.0119} \ x \ H^{0.5234} \ x \ r^{0.1388}) \ x \ 1/v^{0.229}} \quad \text{---- (1)}$$

Subject to the boundary conditions [12]

$v_{min} \leq v \leq v_{max}; \qquad f_{min} \leq f \leq f_{min}$

$d_{min} \leq d \leq d_{min}; \qquad r_{min} \leq r \leq r_{min}$

Table no 1. Boundary conditions of plane turning

Cutting parameter	Range	
	Min	Max
Cutting Speed v in m/min	30	90
	90	180
Feed f in mm/rev	0.2	0.4
	0.4	0.8
Depth of cut d in mm	0.5	1.0
	1.0	1.5
Tool nose radius r in mm	0.4	0.8
	0.8	1.2
Material hardness constant H in BHN	125	

Where v is the cutting speed (m/min), f is the feed rate (mm/rev), d is the depth of cut (mm), r is the nose radius of the tool (mm) and H is hardness constant of the material. Based on the above mentioned different machining conditions sixteen combinations of different operating conditions were indentified for this study. They are:

Table no 2 Ranges of cutting parameters

S. No.	v m/min	f mm/rev	d mm	r mm
1	30-90	0.2-0.4	0.5-1.0	0.4-0.8
2	30-90	0.2-0.4	0.5-1.0	0.8-1.2
3	30-90	0.4-0.8	0.5-1.0	0.4-0.8
4	30-90	0.4-0.8	0.5-1.0	0.8-1.2
5	30-90	0.2-0.4	1.0-1.5	0.4-0.8
6	30-90	0.2-0.4	1.0-1.5	0.8-1.2
7	30-90	0.4-0.8	1.0-1.5	0.4-0.8
8	30-90	0.4-0.8	1.0-1.5	0.8-1.2
9	90-180	0.2-0.4	0.5-1.0	0.4-0.8
10	90-180	0.2-0.4	0.5-1.0	0.8-1.2
11	90-180	0.4-0.8	0.5-1.0	0.4-0.8
12	90-180	0.4-0.8	0.5-1.0	0.8-1.2
13	90-180	0.2-0.4	1.0-1.5	0.4-0.8

S. No.	v m/min	f mm/rev	d mm	r mm
14	90-180	0.2-0.4	1.0-1.5	0.8-1.2
15	90-180	0.4-0.8	1.0-1.5	0.4-0.8
16	90-180	0.4-0.8	1.0-1.5	0.8-1.2

III. QUANTUM INSPIRED EVOLUTIONARY ALGORITHM

Quantum Inspired Evolutionary Algorithm is essentially a stochastic population based evolutionary algorithm that exploits some principles of quantum mechanics, such as Q-bits, superposition, quantum gates and quantum measurement [12].In conventional EAs, encoding the solutions onto chromosomes uses many different representations, which may be generally grouped into three classes: symbolic, binary, and numeric. In contrast, a Quantum Inspired Evolutionary Algorithm uses novel probabilistic representation called as Q-bit. Q – bit is a smallest unit of information that can be in superposition of basis states in a quantum system. Q- Bits are generally represented by a vector in Hilbert space with |0> and |1> as basis states. The Q-bit can be represented as:

$$|\psi> = \alpha|0> + \beta|1> \quad …..(2)$$

Where $|\alpha|^2$ and $|\beta|^2$ are the probability amplitudes of the Q-bit that may exist in state '0' or in state '1' so that it satisfies the normal condition

$$|\alpha|^2 + |\beta|^2 =1 \quad …… (3)$$

Quantum Inspired Evolutionary Algorithm uses a better characteristic of diversity than classical approaches, since it can represent superposition of states. Convergence is also achieved with such representation. As a Q-bit tends towards 1 or 0 during the process of probabilistic observation, the Q-bit converges to a single state and the property of diversity disappears gradually. That is, the Q-bit representation is able to possess the two characteristics of exploration and exploitation, simultaneously. The basic structure of Quantum Inspired Evolutionary Algorithm [13] is presented below:

begin

$t \longleftarrow 0$

initialize $Q(t)$

make $P(t)$ by observing $Q(t)$ states

evaluate $P(t)$

store the best solution among $P(t)$

While (**not** termination – condition) **do**

begin

$t \qquad t + 1 \longleftarrow$

make $P(t)$ by observing $Q(t$-$1)$ states

evaluate $P(t)$

update $Q(t)$ using quantum gates $U(t)$

store the best solution among $P(t)$

end

end

Pseudo code of QIEA

Initialize: Initialize the population Q_{ij}, where i = 1, 2... n, j = 1, 2..q , and n, q are population size and number of parameters respectively.   Assign equal probabilities to α and β of each Q-bit, so that normal condition $|\alpha|^2 + |\beta|^2 = 1$, is satisfied. And set the generation number to 0.

Observe: Observe all the Q-bits. That is, if $|\beta_i|^2 >$ rand, where rand Є [0, 1], then, the observed state would be '1', else, the observed state would be '0'. Decode the binary bits and if necessary employ a repair algorithm to correct boundary violations.

Evaluate: Evaluate the fitness.

Store: Store the best result of generation 0, as f (b). Increment the generation by one and repeat observe and evaluate processes, and store the best result as f (x).

U*pdate***:** Compare each Q-bit of all the parameters pertaining to the best solutions of f (b) and f (x_1). Based on the quantum rotation gate lookup Table 3

and by employing the equation (4), and update the Q-bits. Now repeat *observe, evaluate,* and *update* processes until requirements are met.

$$\begin{bmatrix} \alpha'i \\ \beta'i \end{bmatrix} = \begin{bmatrix} \cos\theta i & -\sin\theta i \\ \sin\theta i & \cos\theta i \end{bmatrix} \begin{bmatrix} \alpha i \\ \beta i \end{bmatrix} \quad .... (4).$$

Once we determine the number of Q-bits per variable, i.e., in this case of plane turning, the cutting speed v is a parameter which varies from 30 to180 m/min, requires eight Q-bits, but for all the other cutting parameters such as feed rate f, depth of cut d, tool nose radius r, we need only four Q-bits. Randomly generate population of parameters, in this case population size is 20, and assign equal probabilities to α and β of each Q-bit of every parameter and conduct probabilistic measurement for observed states of Q-bits, by generating a random number and comparing it with $|\beta|^2$. If rand $> |\beta|^2$ consider the Q-bit as 1 otherwise as 0. Now, boundary violations are checked and repaired, if necessary, by using a repair algorithm.

Table no. 3. Quantum rotation gate lookup table [14]

x_i	b_i	$f(x)$ $\geq$ $f(b)$	$\Delta\theta$	$S(\alpha_i \ \beta_i)$			
				$\alpha_i \ \beta_i$ >0	$\alpha_i \ \beta_i$ <0	$\alpha_i =$ 0	β_i $=0$
0	0	F	0	0	0	0	0
0	0	T	0	0	0	0	0
0	1	F	0	0	0	0	0
0	1	T	0.05 Π	-1	+1	±1	0
1	0	F	0.01 Π	-1	+1	±1	0
1	0	T	0.025 Π	+1	-1	0	±1
1	1	F	0.005 Π	+1	-1	0	±1
1	1	T	0.025Π	+1	-1	0	±1

Now, evaluate the fitness and *store* the best solution among the twenty solutions of generation 0, f (b). Now repeat *observe, repair, evaluate* processes and *store* the best solution among the twenty solutions of generation 1, f (x). Now compare the corresponding Q-bits of all the parameters of best solution f (b) and

best solution f(x), to update Q-bits, by determining rotation angle using quantum rotation gate Table no. 3 and employing equation (4). Here such iterative process of observing, repairing, evaluating and updating is continued till maximum number of cycle is not met, which is sixty in this study.

Experiments on Quantum Inspired Evolutionary Algorithm: The experiments are conducted on a Laptop machine equipped with the processor Intel Core 2 Duo, 4GB RAM and 150 GB HDD. The software is developed in MATLAB 7.0. The program parameters of Quantum Inspired Evolutionary Algorithm (QIEA) are: population size is 20 and maximum generation number 60 and the number of independent simulation runs are 30. The program parameters of Differential Evolution Algorithm (DEA) are: Population size 20, maximum generation number 60, and cross over rate 0.9, mutation rate 0.8, and the number of independent runs are 30.

IV. RESULTS AND DISCUSSION

Quantum Inspired Evolutionary Algorithm is applied on the machining model, referred in section II, for minimum average surface roughness of plane turning, to determine optimal cutting parameters such as cutting speed, feed rate, depth of cut and tool nose radius, for corresponding sixteen operating conditions referred in Table 2.

Table no 4. QIEA determined surface roughness

S. No.	RCGA [6]	DEA [7]	QIEA
1	0.857260	0.8035080	**0.803017**
2	0.928069	0.8846534	**0.884113**
3	1.786641	1.6292232	**1.628231**
4	1.880024	1.7937567	**1.792665**
5	0.851825	0.8101631	**0.809668**
6	0.928080	0.8919806	**0.891436**
7	1.836571	1.6427173	**1.641714**
8	1.878548	1.8086136	**1.807512**
9	0.766149	0.685574	**0.685155**
10	0.817921	0.7548094	**0.754348**

S. No.	RCGA [6]	DEA [7]	QIEA
11	1.512453	1.3900957	**1.389251**
12	1.690667	1.5304801	**1.529550**
13	0.738389	0.6912523	**0.690832**
14	0.810552	0.7610611	**0.760596**
15	1.504253	1.4016093	**1.400753**
16	1.687547	1.5431564	**1.542211**

The average surface roughness as determined by QIEA algorithm is presented in Table no 4. The average surface roughness predicted by QIEA is compared with the estimations of real coded genetic algorithm (RCGA) and differential evolution algorithm (DEA). The results demonstrates that, QIEA has outperformed RCGA and improved results over DEA in achieving better surface quality, for plane turning, in every given operating environment. The below shown figure no 1, depicts the improved performance of Quantum Inspired Evolutionary Algorithm over Real Coded Genetic Algorithm and Differential Evolution Algorithm

Table no 5. QIEA determined optimal cutting speed

| S. No.	v m/min RCGA [6]	v m/min DEA [7]	v m/min QIEA
1	82.717673	89.9996112	90.000000
2	86.577654	89.9912760	90.000000
3	62.817164	89.9975154	90.000000
4	87.211219	89.9993688	90.000000
5	85.125584	89.9993875	90.000000
6	89.252907	89.9994307	90.000000
7	87.040925	89.9993405	90.000000
8	85.768303	89.9999034	90.000000
9	173.232215	180.000000	180.000000
10	160.781579	180.000000	180.000000
11	138.670919	180.000000	180.000000
12	156.988433	179.999994	180.000000
13	147.707450	179.999989	180.000000
14	172.982269	179.999997	180.000000
15	179.980773	179.999991	180.000000
16	158.174993	180.000000	180.000000
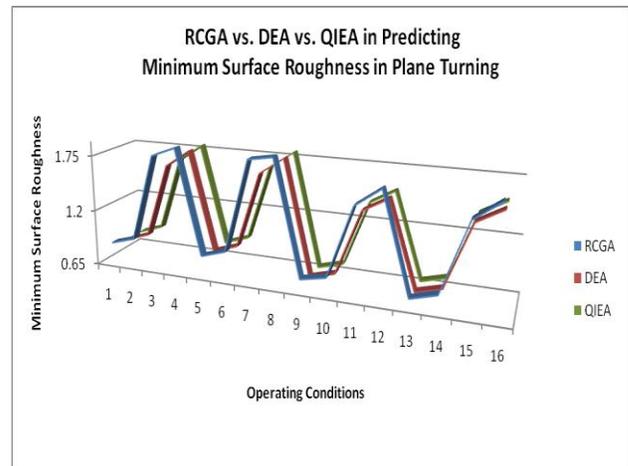
Figure no1. RCGA vs DEA vs QIEA in predicting minimum surface roughness in plane turning

Table no 6. QIEA determined optimal feed rate

S. No.	f mm/rev RCGA [6]	f mm/rev DEA [7]	f mm/rev QIEA
1	0.200012	0.200053	0.200000
2	0.204248	0.200034	0.200000
3	0.401709	0.400031	0.400000
4	0.402454	0.400060	0.400000
5	0.205896	0.200000	0.200000

S. No.	f mm/rev RCGA [6]	f mm/rev DEA [7]	f mm/rev QIEA
6	0.206922	0.200079	0.200000
7	0.432032	0.400013	0.400000
8	0.409876	0.400005	0.400000
9	0.200159	0.200000	0.200000
10	0.207935	0.200000	0.200000
11	0.405750	0.400000	0.400000
12	0.420301	0.400000	0.40000
13	0.203143	0.200000	0.200000
14	0.202057	0.200000	0.200000
15	0.408081	0.400000	0.400000
16	0.419410	0.400000	0.400000

S. No.	d mm RCGA [6]	d mm DEA [7]	d mm QIEA
5	1.168371	1.0003244	1.000000
6	1.005707	1.0033508	1.000000
7	1.272668	1.0003603	1.000000
8	1.151006	1.0002004	1.000000
9	0.898267	0.000003	0.500000
10	0.696722	0.5000003	0.500000
11	0.650456	0.5000001	0.500000
12	0.697592	0.5000000	0.500000
13	1.012772	0.5000002	1.000000
14	1.230903	1.0000006	1.000000
15	1.375683	1.0000008	1.000000
16	1.069231	1.0000004	1.000000

Reaffirming the findings [5, 6, 7], the important cutting parameters, cutting speed and feed rate, that are significantly affect the surface roughness are tabulated in Table no 5 and 6, as predicted using QIEA algorithm. These results confirm that, at higher cutting speed and at lower feed rate minimum surface roughness can be achieved.

The other two cutting parameters, depth of cut and tool nose radius are presented in the Table no 7 and 8. At minimal settings of these two parameters the improved surface finish is achieved by the algorithm. With help of these results of QIEA estimations of plane turning, it is observed that at higher cutting speeds and at lower feed rate, and at minimal settings of other two parameters, best surface roughness can be achieved.

Table no 7. QIEA determined optimal depth

S. No.	d mm RCGA [6]	d mm DEA [7]	d mm QIEA
1	0.720191	0.5007623	0.500000
2	0.913816	0.5000292	0.500000
3	0.726173	0.5000123	0.500000
4	0.817698	0.5001041	0.500000

Table no 8. QIEA determined nose radius

S. No.	r mm RCGA [6]	r mm DEA [7]	r mm QIEA
1	0.537565	0.4001902	0.400000
2	0.862380	0.8004704	0.800000
3	0.403186	0.4000622	0.400000

S. No.	r mm RCGA [6]	r mm DEA [7]	r mm QIEA
4	0.976519	0.8000915	0.800000
5	0.417408	0.4000517	0.400000
6	0.817481	0.8000086	0.800000
7	0.470266	0.4001726	0.400000
8	0.802100	0.8000237	0.800000
9	0.790576	0.4000000	0.400000
10	0.864846	0.8000000	0.800000
11	0.420508	0.4000000	0.400000
12	0.883425	0.8000000	0.800000
13	0.413575	0.4000000	0.400000
14	1.074825	0.8000000	0.800000
15	0.559014	0.4000000	0.400000
16	0.864284	0.8000000	0.800000
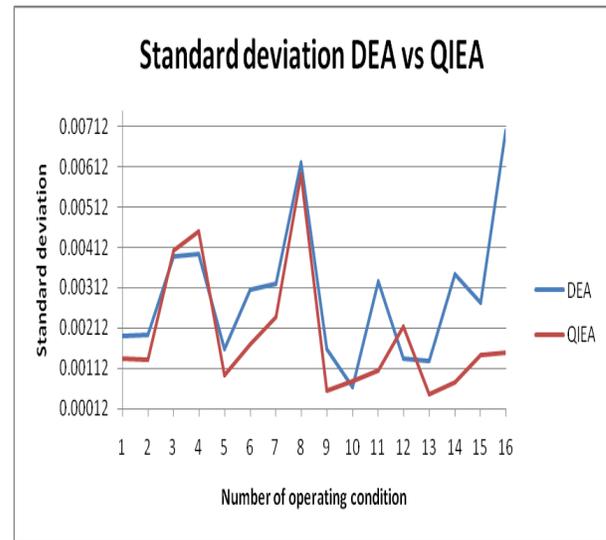
Figure no2. Standard deviation of DEA vs.QIEA

Table no 9. Standard deviation - DEA vs. QIEA

Since the performance details of the Real Coded Genetic Algorithm (RCGA) and Differential Evolution Algorithm on this plane turning optimization of problem, in terms of its standard deviation, mean of worst and mean of best is not reported [], in this work an effort is made to bring out such results for DEA and compare the same with Quantum Inspired Evolutionary Algorithm   The thirty independent experiments carried out on Quantum Inspired Evolutionary Algorithm are compared with the thirty independent runs carried out on Differential Evolution Algorithm (DEA) for the same problem in order to ascertain the performance of QIEA.

S. No.	DEA Standard deviation	QIEA Standard deviation
1	0.0019217	0.00136438
2	0.0019474	0.00134163
3	0.0038966	0.00403858
4	0.0039486	0.00450954
5	0.0015856	0.00096127
6	0.0030712	0.00170577
7	0.003215	0.00238215
8	0.0062272	0.00591869
9	0.0016095	0.00057922
10	0.0006445	0.00081122

S. No.	DEA Standard deviation	QIEA Standard deviation
11	0.003263	0.00105965
12	0.00137	0.00216368
13	0.00131	0.00047915
14	0.003457	0.00078677
15	0.002739	0.00143839
16	0.007009	0.00150446

The comparisons are drawn between QIEA and DEA by presenting the contrasts of different statistical parameters i.e., standard deviation, mean of worst, mean of best. The results of such independent experimental runs of DEA and QIEA are presented in Table no 9 and 10. The following figure no 2, depicts the comparison of standard deviation between Quantum Inspired Evolutionary Algorithm and Differential Evolution Algorithm.

Table no 10. Mean of worst and best DEA vs. QIEA

S. No	Mean Worst		Mean Best	
	DEA	QIEA	DEA	QIEA
1	0.8109615	0.8443	0.8035781	0.803017
2	0.8922561	0.972828	0.8849536	0.884113
3	1.6443363	1.78902	1.6293654	1.62831
4	1.8091721	2.04418	1.7943654	1.79266
5	0.8165556	0.946924	0.8103791	0.809668
6	0.9043026	0.93703	0.8921801	0.891436
7	1.6556791	1.97014	1.6431554	1.641715

S. No	Mean Worst		Mean Best	
	DEA	QIEA	DEA	QIEA
8	1.8335982	1.92121	1.809018	1.807513
9	0.6929077	0.733505	0.685813	0.685152
10	0.7583466	0.781242	0.755081	0.754341
11	1.404966	1.60212	1.39058	1.38925
12	1.537652	2.03124	1.531031	1.52955
13	0.698588	0.69718	0.691327	0.690831
14	0.780506	0.787425	0.76114	0.760592
15	1.416483	1.56292	1.401761	1.400831
16	1.582584	1.63311	1.543315	1.54221

V. CONCLUSION

Quantum Inspired Evolutionary Algorithm which essentially exploits some quantum mechanics principles such as Q- bit, superposition, quantum measurement, quantum gate is successfully implemented in MTALAB 7.0 environment. The QIEA is applied on a machining model of turning to estimate optimal cutting parameters, such as cutting speed, depth of cut, feed rate, nose radius of the tool, for sixteen different operating conditions, to achieve improved surface finish. Estimations of Quantum Inspired Evolutionary Algorithm, suggest that the surface roughness in turning operation is significantly affected by cutting speed and feed rate. It is observed that low feed rate would give better surface quality. These observations are in concurrence to the reported findings [6]. The results obtained by QIEA are better than RCGA and are comparable to those of DEA. QIEA is found to be computationally efficient and suitable for machining optimization. The statistical results of independent experiments reveal that Quantum Inspired Evolutionary Algorithm is comparably effective in terms of stability when compared to Differential Evolution Algorithm. Therefore, Quantum Inspired Evolutionary Algorithm is a promising heuristic for intelligent manufacturing.

REFERENCES

[1] A. Bhattacharya, R. Faria-Gonzalez and I. Ham, "Regression analysis for predicting surface finish and its application in the determination of optimum machining conditions", *ASME Journal of Engineering for Industry, 4, pp. 711–714*, 1970.

[2] Chandrasekaran, M. Muralidhar, C. Murali Krishna, U. S. Dixit, "Application of soft computing techniques in machining Performance prediction and optimization: A literature review," *Int J Adv Manuf Technol, 46:445–464.*, 2010.

[3] Suresh, P.V.S., Venkateswara P. Rao, S. G. Deshmukh,, "A Genetic Algorithmic Approach for Optimization of Surface Roughness Prediction Model", *International Journal of Machine Tools & Manufacture, 42, p: 675–680,* 2002.

[4] Franci Cus, Joze Balic,, "*Optimization of cutting process by GA approach*", Robotics and Computer Integrated manufacturing 19, pp.113-121, 2003.

[5] T. Srikanth, Dr V. kamala, "*Experimental determination of optimal speeds for alloy steels in plane turning*", Proceedings of the *9th Biennial ASME Conference on Engineering Systems Design and Analysis*,ESDA2008, Haifa, Israel., 2008.

[6] T. Srikanth, Dr V. kamala, "*A Real Coded Genetic Algorithm for Optimization of Cutting Parameters in Turning*" International Journal of Computer Science and Network Security, VOL.8 No.6, June 2008, pp. 189-193, 2008.

[7] R.S.S. Prasanth, K. Hans Raj, "Application of differential evolution algorithm for optimizing orthogonal cutting". *Proceedings of International conference on Systemics, Cybernetics, and Informatics. Pp 122 – 126.* 2011.

[8] Narayanan, A., Moore, M.: Quantum-inspired genetic algorithms. *In: Proc. CEC, pp. 61–66*, 1996.

[9] Ashish Mani and C. Patvardhan., "An adaptive quantum evolutionary algorithm for Engineering optimization problems". *International Journal of Computer Applications, 1(22):43–48.*, 2010

[10] Ashish Mani and C. Patvardhan., "Solving Ceramic Grinding Optimization Problem by Adaptive Quantum Evolutionary Algorithm", *Intelligent Systems, Modeling and Simulation, International Conference on, 0:43–48*, 2010.

[11] Gexiang Zhang,,"Quantum-inspired evolutionary algorithms: a survey and empirical study", *J Heuristics, 17: 303–351*, 2011.

[12] K.H. Han and J.H. Kim., "Genetic quantum algorithm and its application to combinatorial optimization problem". *In Proceedings of the 2000 Congress on Evolutionary computation, volume 2, pages 1354–1360*. Citeseer, 2000.

[13] *K.H. Han and J.H. Kim,* "Quantum-inspired evolutionary algorithm for a class of combinatorial optimization" *IEEE transactions on evolutionary computation, 6(6):580–593,* 2002.

[14] Kuk-Hyun Han*., "Quantum-inspired Evolutionary Algorithm. PhD thesis",* Korea Advanced Institute of Science and Technology (KAIST), 2003

An Industrial Energy Auditing: Basic Approach

Mr. Nilesh R. Kumbhar[1], Mr. Rahul R. Joshi[2]

(*Dept. Of Mechanical Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur.Shivaji University, Kolhapur, India)

(** Dept. Of Mechanical Engineering, Dr. J. J. Magdum College of Engineerin, Jaysingpur.Shivaji University, Kolhapur, India)

ABSTRACT:

Growing concerns arise about energy consumption and its adverse environmental impact in recent years in India, which cause manufactures to establish energy management groups. The energy auditing is the key to successful running of an industry with saving energy & contributing toward preserving national recourses of energy. "Managing energy is not a just technical Challenge but one of how to best implement those technical Challenges within economic limits, and with a minimum of disruptions. In this paper importance of energy auditing and process of energy audit is discussed.

Keywords: **Energy Audit, Energy Audit Report, ECOs, Energy Saving.**

I. INTRODUCTION

This Saving Money on energy bills is attractive to Business, industries and individuals alike customers whose energy bill use up a large part of their income and especially those customers whose energy bills represent substantial fraction of their companies operating cost, have strong motivation to initiate and continue on an ongoing to energy cost control program. No cost or very low cost operational changes can often save a customer or an industry 10-20% on utility bills Capital Cost Programs with pay back times of two years or less can often save an additional 20-30%[1].

The energy auditing is one of the first task to be Promoted in the accomplishment of an effective energy cost control Program .An energy audit consist of a detailed examination of a how facility uses energy ,what the facility pays for that energy ,and a finally, a recommended program for changes in operating   practices or energy consuming equipment that will cost effectively saves bucks on energy bills.

With new technology and alternative energy resources now available, this country could possibly reduce its energy consumption by 50%. If there were no barriers to implementation [2] but off course there are barriers mostly economical.

Energy auditing is an official method of finding out the ECO's. Ii is the official survey / study of the energy consumption / processing / supply aspects related with of industry or organization. Purpose of energy auditing is to recommend steps to be taken by Management for improving the energy efficiency, reduce energy cost and saving the money on the energy bills.

II. METHODS OF ENERGY AUDITING:

Energy audits can be carried outs in different ways. Depending on time span invested auditing can be classified in as:

i)   *Walk Through Audit*
ii) *Intermediate Audit*
iii) *Detailed / Comprehensive Audi*

i)   *Walk Through Audit*

This is simple kind of energy, it carries rapid survey of plant. During rapid walk survey main focus is on the energy input, spots of energy wastages and ECO's. Data about plant is collected in such a way that, data should be utilized for next detailed audits. Usually audit is carried out at two periods viz. During off period & during working shifts

Generally this kind of audit is carried out for three days to one week. As the time span required is short cost involve in auditing is also less.

ii) *Intermediate Audit*

This kind of audit is conducted for detailed survey and measurement of systems compare with walk through audit. Major focus is made on energy loses measure and quantification to analyze energy efficiency of system. Generally low tech recommendations are preferred with first preference is given for

-Switching off lights and fans when not required.
-Placing of automatic thermostat to control temperature of water heaters etc
-Spotting out golden ECO's which involves higher energy wastage cost.

This type of audit is carried out for one week to one week; time span required is more so the cost associated with audit is also more compare with walk through audit.

iii) Detailed / Comprehensive Audit

This is exhaustive audit than the previous types of audit. Detailed survey of systems as well as subsystems of an industry is done. Energy consumption of all subsystems and systems is compared with targeted energy consumption. This kind of audit also identifies the consumption of secondary energy like electricity, steam, gases etc. Modernization and changes in major retrofitting as suggested if required.

III. BASIC COMPONENTS OF EVERY AUDITING

The Energy Audit Process starts by collecting information about facilities Operation and its past record of utility bills. This data is then analysed to get Picture of how the Facility uses and possibly wastes energy, as well as to help the auditor learn that areas to examine to reduce energy cost. Specific changes called Energy Conversion Opportunities (ECO) are identified and evaluated to determine their benefits and their cost effectiveness. These ECOs are accessed in terms of their costs & benefits and economic comparison is made to rank various ECOs. Finally an action plan is created whether certain ECOs are selected for implementation and the actual process of energy saving & saving money begins [4],[5].

i) Auditor's tool box: To obtain the best information from a successful energy cost control program the auditor must make some measurement during audit visit.

ii) Preparation for audit visit: Some preliminary work must be done before the auditor makes actual energy audit. To a facility some parameters that should be needed are: energy use data, energy rate schedule, physical & operational data for facility that will consist of geographical location, whether data, facility layout, operation house, equipment list. One more important part of energy audit is safety of energy auditor & audit team. The audit person & audit team must be thoroughly briefed on safety equipments & processes.

iii) Conducting the audit: Once the information on energy bills, faculty equipments and facility operations has been obtained, the audit equipment can be gathered up and actual visit is to be started. Following are some important steps in audit:

iv) Introductory meeting audit team should meet facility manager & maintenance manager to brief about purpose of audit

v) Audit interview getting correct information on facility equipment and operation is important, if the audit is going to most successful in identifying ways to save money on energy bills. Auditor must interview with floor supervisor and equipment operator to understand building and process problems.

vi) Walk through audit a walk through tour of facility or plant should be arranged by facility/ plant manager and should be arranged to the auditor or audit team can see major operational and equipment features of facility. During walk through audit data regarding ECOs should be gathered by looking at : lighting, HVAC system, electrical motors, water heaters, waste heat sources, peak equipment loads and other energy consuming equipments.

vii) Post audit analysis after visit data collected should be examined , organized and reviewed for completeness ant thing missing data items should be obtained form facility of re-visit.

viii) The energy audit report: Next step in energy auditing process is to prepare a report which details the final result and recommendation. An industrial audit report is more likely to have a detailed explanation of ECOs and benefit cost analysis. The report should begin with an executive summary that provide owners/ manager of facility with brief synopsis of total saving available and the highlights of each ECOs

ix) Energy action Plan: The last step in audit process is to recommend an action plan for facility. The energy action plan list the ECOs which should implement first and suggest an over all implementation schedule , often one or more of the recommended ECOs should provide an immediate or very short period pay back, so saving from that Eco or those ECOs can be used to generate capital to pay for implementing other ECO

IV. CONCLUSIONS

Energy audit is an effective tool in identifying and perusing a comprehensive energy management program. A care full audit of any type will give the industry a plan with which it can effectively manage the industrial energy system at minimum energy cost. This approach could be useful for an industry in combating essential energy cost and also raps several other benefits like improved production, better quality, higher profit and most important satisfaction of heading towards contributing in world energy saving.

REFERENCES:

[1]   William H. Mashburan, P.E., CEM " Effective Energy Management"

[2]   Barney L Capehar and Mark B. Spillter " Energy Auditing",

[3]   Hersey, Paul and Kenneth " management of organizational behavior"   1970

[4] Insruction for energy auditor, *Vol I & II U.S. Dept of Energy, Sept 1978*

[5]   Energy conversion guide for industry & Commerce, *hand book & supplement*

AN EFFICIENT WAY OF RETRIEVAL DATA BY TRACKING ATTACKERS

G.Sindhu*, Mrs.R.Kalaiselvi**

*II M.E CSE, Sri Shakthi Institute Of Engineering and Technology, Anna University, Coimbatore
**Asst.prof CSE, Sri Shakthi Institute Of Engineering and Technology, Anna University, Coimbatore

ABSTRACT:

A major threat to the internet is that the Distributed Denial- of- Service (DDoS) attacks. There is no efficient way to traceback the attackers because of memoryless feature of routers. In this paper, trace back of the attackers in a wireless networks are efficiently identified and also to protect the data from the attackers using entropy variations. In the existing system, some approaches have been suggested to identify the attackers such as probabilistic Packet Marking (PPM), Deterministic Packet Marking (DPM). These two approaches are not efficient because it requires injecting marks into individual packets in order to trace back the attackers. In PPM, it can only operate in a local range of internet. In DPM, it requires all the internet routers to be updated for packet marking. Scalability is also a big problem in both PPM and DPM. In order to overcome the above drawbacks, a method based on Entropy Variation is proposed which is a measure changes of randomness of flows at a router for a given interval in a large scale attack network. This method is used to identify the attackers efficiently and supports a large scalability.

Index terms - DDOS, IPtraceback, Entropy Variation.

I.INTRODUCTION

A DoS(denial of service) attack is a malicious attempt by a single person or a group of people to cause the victim, site, or router to deny service to its customers. When this attempt derives from a single host of the network, it constitutes a DoS attack. On the other hand, it is also possible that a lot of malicious hosts coordinate to flood the victim with an abundance of attack packets, so that the attack takes place simultaneously from multiple points. This type of attack is called a Distributed *DDoS,* or DDoS attack. A "denial-of-service" attack is characterized by an explicit attempt by attackers to prevent legitimate users of a service from using that service. There are two general forms of DoS attacks: those that crash services and those that flood services. Attacks can be directed at any network device, including attacks on routing devices and web, electronic mail, or Domain Name System servers. The main reason is that the network security community does not have efficient and effective trace back methods to locate the attackers in a wireless network as it is easy for attackers by taking the advantages of vulnerabilities of the world wide web [11]. In this Traceback of DDoS Attacks Using Entropy Variations is used to find out the

Distributed Denial-of-Service (DDoS) attacks are a critical threat to the Internet and user's location with the help of Entropy Variation Mechanisms against DPM and PPM.

Traceback of DDoS Attacks is random. Therefore, a Entropy Variation Mechanisms against DPM and PPM should empower a router with the ability to determine whether it should move and where it should move to such that the movement can enhance Attacks quality without depleting scarce resources or significantly compromising coverage and network connectivity. It is the movement of the routers are purposeful. It is important to have an efficient Traceback of DDoS Attacks scheme to ensure that the sensor router mobility is exploited in the best possible way. At the same time the mobility management strategy should avoid inefficient usage of scarce resources, such as energy and network bandwidth. Vulnerable hosts are those that are either running no antivirus or out-of-date antivirus software. These are exploited by the attackers who use the vulnerability to gain access to these hosts. The next step for the attacker is to install new programs on the compromised hosts of the attack network. The hosts running these attack tools are known as zombies and they can be used to carry out any attack under the control of the attacker..IP trace back methods should be independent of packet pollution and various attack patterns. Because of the vulnerability, the original attackers cannot be found. An ad hoc network is a collection of mobile hosts forming a temporary network. The transmission of a mobile host is received by all hosts within its transmission range due to broadcast of wireless communication and Omni directional antennae.

In the existing system, there are two major methods for IP Trace back, Probabilistic Packet Marking (PPM) [1], [2], [3] and Deterministic Packet Marking (DPM) [4], [5]. The DPM strategy requires all the routers to be updated for packet marking. Hence, the scalability of DPM is a huge problem. Moreover, the DPM mechanism poses an extra ordinary challenge on storage for packet logging for routers. Both PPM and DPM are vulnerable to hacking, which is referred to as packet pollution. IP trace back methods should be independent of packet pollution and various attack patterns. Therefore, an entropy variation mechanism empowers a router with the ability to determine whether it should move and where it should move to such that the movement can enhance attack quality. It is important to have an efficient traceback scheme to ensure that the sensor router mobility is exploited. At the same time the mobility management

strategy should avoid inefficient usage of scarce resources such as energy and network bandwidth.

II. AN OVERVIEW OF RELATED WORK

It is obvious that to trace back the attackers is essential in solving the DDOS attacks. In general, the trace back strategies are based on packet marking which include PPM and DPM. It is an extra ordinary challenge to traceback the source of Distributed Denial-of-Service (DDoS) attacks in the Internet. In DDoS attacks, attackers generate a huge amount of requests to victims through compromised computers (zombies), with the aim of denying normal service or degrading of the quality of services. The key reason behind this phenomena is that the network security community does not have effective and efficient traceback methods to locate attackers as it is easy for attackers to disguise themselves by taking advantages of the vulnerabilities of the World Wide Web, such as the dynamic, stateless, and anonymous nature of the Internet. IP traceback means the capability of identifying the actual source of any packet sent across the Internet. Because of the vulnerability of the original design of the Internet, the actual hackers may not be able to find at present. In fact, IP traceback schemes are considered successful if they can identify the zombies from which the DDoS attack packets entered the Internet.

A number of IP traceback approaches have been suggested to identify attackers and there are two major methods for IP traceback, the probabilistic packet marking (PPM) and the deterministic packet marking (DPM). Both of these strategies require routers to inject marks into individual packets. Moreover, the PPM strategy can only operate in a local range of the Internet (ISP network), where the defender has the authority to manage. However, this kind of ISP networks is generally quite small, and we cannot traceback to the attack sources located out of the ISP network. The DPM strategy requires all the Internet routers to be updated for packet marking. However, with only 25 spare bits available in as IP packet, the scalability of DPM is a huge problem. Moreover, the DPM mechanism poses an extraordinary challenge on storage for packet logging for routers . Therefore, it is infeasible in practice at present. Further, both PPM and DPM are vulnerable to hacking, which is referred to as packet pollution.

The PPM mechanism tries to mark packets with the router's IP address information by probability on the local router, and the victim can reconstruct the paths that the attack packets went through. The PPM method is vulnerable to attackers, as stated in [7], as attackers can send spoofed marking information to the victim to mislead the victim. The accuracy of PPM is another problem because the marked messages by the routers who are closer to the leaves could be overwritten by the downstream routers on the attack tree .At the same time, most of the PPM algorithms suffer from the storage space problem to store large amount of marked packets for reconstructing the attack tree [1], [3].

Based on the PPM mechanism, Law et al. tried to trace back the attackers using traffic rates of packets, which were targeted on the victim [2]. Both of these strategies require routers to inject marks into individual packets.PPM strategy can only operate in a local range of the Internet, where the defender has the authority to manage. However, this kind of ISP Networks is generally small and we cannot trace back to the attack sources located of the ISP Network. The model bears a very strong assumption: the traffic pattern has to obey the Poisson distribution, which is not always true in the Internet. Moreover, it inherits the disadvantages of the PPM mechanism: large amount of marked packets are expected to reconstruct the attack diagram, centralized processing on the victim, and it is easy be fooled by attackers using packet pollution.

The deterministic packet marking mechanism tries to mark the spare space of a packet with the packet's initial router's information, e.g., IP address. Therefore, the receiver can identify the source location of the packets once it has sufficient information of the marks. The major problem of DPM is that it involves modifications of the current routing software, and it may require very large amount of marks for packet reconstruction. Moreover, similar to PPM, the DPM mechanism cannot avoid pollution from attackers.

Savage et al. [3] first introduced the probability-based packet marking method, router appending, which appends each router's address to the end of the packet as it travels from the attack source to the victim. Obviously, it is infeasible when the path is long or there is insufficient unused space in the original packet.

Snoeren et al. proposed a method by logging packets or digests of packets at routers [9]. The packets are digested using bloom filter at all the routers. Based on these logged information, the victim can trace back the leaves on an attack tree. The methods can even trace back a single packet. However, it also places a significant strain on the storage capability of intermediate routers.

III.   ENTROPY   VARIATION   BASED TRACEBACK   MECHANISM

Entropy variation is a measure of randomness flow of the routers at a given interval of time. The parameters to identify the attackers are time between the two routers in which the data was sent and delay for the overall routers. This mechanism comprises of two algorithms to traceback the attackers and to retrieve the original data.

The flow monitoring algorithm monitors the flow of each and every router. The packets that are passing through the routers are categorized into flows. A flow is defined by a pair-the upstream router where the packets came from and the destination address of the packet. A

router knows its local topologies such as its upstream router attached to another router in a local area network. In this paper, I is denoted as the set of positive integers, and R as the set of real numbers. A flow on a local router is denoted by $<u_i, d_j, t>$; I, j € I, t €R, where u_i is an upstream router of a local router R_i, d_j is the destination generated at the local area network which is the local flows, and L is used to represent the local flows. All the incoming flows are represented as input flows, and all the flows leaving router R_i are named as output flows. We denote u_i; i 2 I as the immediate upstream routers of the local router R_i, and set U as the set of incoming flows of router R_i. Therefore, U={u_i,i €I} +{L}. We use a set D ={d_i, i € I} to represent the destinations of the packets that are passing through the local router R_i. If v is the victim router, then v € D. Therefore, a flow at a local router can be defined as follows:

$$F_{ij}(u_i, d_j)=\{<u_i, d_j, t>/u_i € U; d_j €D,I, j € I\}$$

The trace back mechanism performs in terms of scalability which is the size of the networks that can be handled, the storage space that need on routers, trace back time and the operation workload. During non attack period, local flow monitoring is done by gathering information from normal network flows progressing the mean and standard variation of flows. Once a DDOS attacks has been confirmed, the victim starts the IP trace back algorithm. In order to make analysis simple and clear some assumptions are made:

1. The changes may occur through network traffic in a very long time interval for non-DDoS attack cases. By breaking the long time interval into seconds, the change of traffic is recognized.

2. The number of attack packets is much higher than that of legitimate flows. For a local router, the number of flows is N and the probability is P(p1,p2,p3,…..pn) is considered. By considering the flows the attackers are identified and the original data is obtained.

IV. PERFORMANCE EVALUATION

A. Simulation model

Consider an ad hoc network in which routers are uniformly distributed in a square area. In the network, sessions are generated between randomly chosen source-destination routers with exponentially distributed inter-arrival time. The source router of the session transmits data packets with the constant rate 1 packet/sec. We developed our simulation model using ns 2.34 simulator. NS-2 simulator allows extracting from a simulation many interesting parameters, like throughput, packet delivery ratio, end-to-end delay.

B. Simulation Results

The following results show the Some parameters like, end to end delay, throughput and packet delivery ratio are analyzed.

Packet delivery ratio:

Data packet delivery ratio can be calculated as the ratio between the number of data packets that are sent by the source and the number of data packets that are received by the sink. This is the amount of successful received bits at the destination routers for the entire simulation period. Packet delivery ratio should be always high for the efficient algorithm or a protocol. The below figure shows the packet delivery ratio was high when compared with the previous methodology.
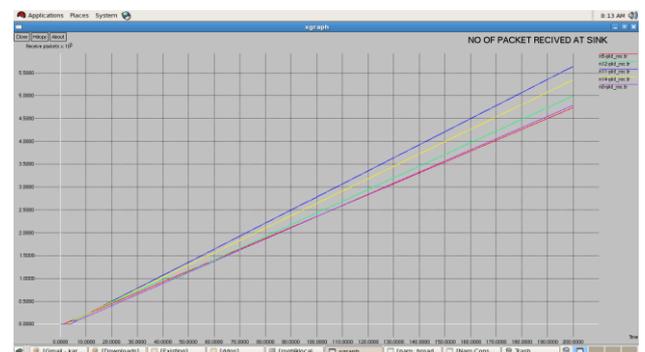

Fig 1: Packet Delivery Ratio

End To End Delay:

End-to-end delay refers to the time taken for a packet to be transmitted across a network from source to destination. End to end delay which includes all possible delays caused by buffering during route discovery time, queuing at the interface queue, retransmission, and processing time. It defines the ratio of interval between the first and the second packets to a total packets delivery.

Fig 2: End to End delay

Throughput:

Throughput is the average rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network router. The throughput is usually measured in bits per second (bit/s or bps), and sometimes in data packets per second or data packets per time slot.


Fig 3: Throughput

Entropy Variation:

By the comparison of the parameters such as the packets received at the sink, End to End delay, Throughput, the actual position of the attackers are identified. The comparison of such parameters are shown in the graph as shown below.
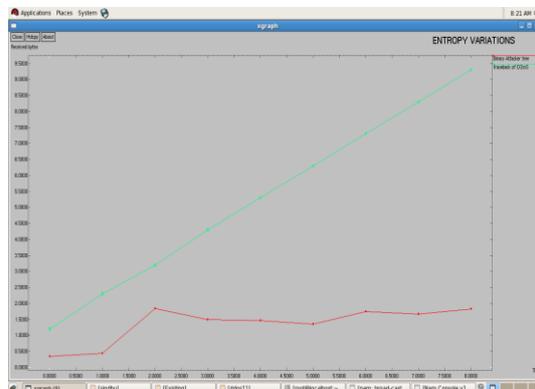
Fig 4:Entropy Variation

In this graph the comparison between the attack and the non attack path is shown. The variation in the attack path is accurately identified and the original data is retrieved.

VI.CONCLUSION

In this paper, the traceback mechanism based on entropy variation is much efficient when compared to the probabilistic packet marking or deterministic packet marking. Because of the vulnerability of the Internet, the packet marking mechanism suffers a number of serious drawbacks: lack of scalability; vulnerability to packet pollution from hacker.

REFERENCES

[1] M. T. Goodrich, "Probabilistic Packet Marking for Large- Scale IP Traceback," IEEE/ACM Trans. Networking, vol. 16, no. 1, pp. 15-24, Feb. 2008.

[2] T. K. T. Law, J. C. S. Lui, and D. K. Y. Yau, "You Can Run, But You Can't Hide: An Effective Statistical Methodology to Traceback DDoS Attackers," IEEE Trans. Parallel and Distributed Systems, vol. 16, no. 9, pp. 799-813, Sept. 2005.

[3] S. Savage, "Network Support for IP Traceback," IEEE/ACM Trans. Networking, vol. 9, no. 3, pp. 226-237, June 2001.

[4] Belenky and N. Ansari, "IP Traceback with Deterministic Packet Marking," IEEE Comm. Letters, vol. 7, no. 4, pp. 162-164, Apr. 2003.

[5] D. Dean, M. Franlin, and A. Stubblefield, "An Algebraic Approach to IP Traceback," ACM Trans. Information and System Security, vol. 5, no. 2, pp. 119-137, May 2006.

[6] G. Jin and J. Yang, "Deterministic Packet Marking Based on Redundant Decomposition for IP Traceback," IEEE Comm. Letters, vol. 10, no. 3, pp. 204-206, Mar. 2006.

[7] K. Park and H. Lee, "On the Effectiveness of Probabilistic Packet Marking for IP Traceback under Denial of Service Attack," Proc. IEEE INFOCOM, 2001.

[8] Gong and K. Sarac, "A More Practical Approach for Single- Packet IP Traceback Using Packet Logging and Marking," IEEE Trans. Parallel and Distributed Systems, vol. 19, no. 10, pp. 1310- 1324, Oct. 2008.

[9] C. Snoeren et al., "Single-Packet IP Traceback," IEEE/ACM Trans. Networking, vol. 10, no. 6, pp. 721-734,Dec.2002.

[10] D. Moore et al., "Inferring Internet Denial-of-Service Activity," ACM Trans. Computer Systems, vol. 24, no. 2, pp.115-139, May 2006.

[11] Patrikakis, M. Masikos, and O. Zouraraki, "Distributed Denial of Service Attacks," The Internet Protocol J., vol. 7, no. 4, pp. 13-35, 2004.

Improving the Security of Cloud Computing using Trusted Computing Technology

P.Senthil[1]
PG Student, Dept of CSE,
CIET, Coimbatore, TN, India.

N.Boopal[2]
Assistant Professor, Dept of CSE,
CIET, Coimbatore, TN, India.

R.Vanathi[3]
PG Student, Dept of CSE,
CIET, Coimbatore, TN, India.

ABSTRACT

Cloud Computing is a collection of computers and servers that are publicly accessible via Internet. It is a significantly new idea that influence the power of internet to process, store and share data from a network of remote servers located anywhere in the world. That is a good way to share a many kinds of distributed resources, but it also makes security problems more complicate and more important for users than before. This paper analyses some security services in cloud computing environment and a method to build a trusted computing environment for cloud computing system by integrating the trusted computing platform into cloud computing system. Trusted Computing Platform (TCP) model can improve the cloud computing security and will not bring much complexity to users. In this model, some important security services including encryption, authentication, integrity and confidentiality are provided in cloud computing system.

Keywords **- Cloud Computing, Trusted Computing, Trusted Computing Platform, TCPA, Trusted Security Services.**

1. INTRODUCTION

Cloud computing provides a large business model that supports pay-for-use, on-demand and economies-of-scale IT services through the Internet. The Internet cloud working as a service factory that built around virtualized data centers. Cloud Computing platforms are dynamically built through the virtualization with provisioned hardware, software, datasets and networks. Cloud computing is an internet based progress and use of computer technology. It provides the way to share distributed resources and services that be in the right place to different organizations. Since cloud computing share distributed resources via the internet in the open environment, thus it makes security problems important for us to develop the cloud computing application. In this paper, we attention to the security requirements in cloud computing environment. It is a method to build a trusted computing environment for cloud computing system by integrating the trusted computing platform into cloud computing system. A model system in which cloud computing is shared with trusted computing platform with trusted platform module. In this model, some important security services, includes authenticated boot, encryption authentication, confidentiality and integrity, are in cloud computing system[1].

Cloud computing technology is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers over a network.Cloud computing is an growing

computing paradigm in which resources of the cloud computing infrastructure are provided as services over the Internet, where a large team of systems are connected in private or public networks, to provide dynamically scalable communications for application, data and file storage. With the arrival of this technology, the cost of computation, application hosting, content storage and delivery is reduced significantly. Cloud computing is a useful approach to experience direct cost benefits and it has the possible to transform a data center from a capital-intensive set up to a variable priced environment[2].

The idea of cloud computing is based on a very fundamental principal of reusability of Information Technology capabilities. The difference that cloud computing brings compare to traditional concepts of "grid computing", "distributed computing", "utility computing", and "autonomic computing" is to extend horizons across organizational boundaries. The Cloud pertains to all the documents or files reserved by servers from remote locations that can be accessed throughout the Internet. Storing data through the Cloud computing makes it easy for all the parties concerned to be able to recover the information they require. Inside the Cloud, users may be able to store and manage their files for personal use, or for other users to be able to utilize it. Cloud computing is a catch-all turn of phrase that covers virtualized operating systems running on virtual hardware on untold numbers of physical servers. It is a computing paradigm where tasks are assigned to a combination of connections, software and services accessed over a network.

Trusted Computing technology was developed and promoted by the Trusted Computing Group(TCG).It is a group of Microsoft, Intel, IBM, HP and AMD which promotes a pattern for a `more secure' Personal Computer. The Trusted Computing Group project is known by a number of names. `Trusted computing' was the innovative one, and is now used by IBM, while Microsoft calls it `trustworthy computing' and the Free Software organization calls it `treacherous computing' which you can pronounce according to taste [16].Trusted Computing Platform Alliance (TCPA),is an initiative started by AMD, HP, IBM, Intel, and Microsoft to implement Trusted Computing.

Trusted computing is a broad word that refers to technology and proposals for resolving computer security problems through hardware enhancements and related software modifications. A number of major hardware manufacturers and software vendors, collectively known as the Trusted Computing Group, are cooperating in this venture and have come up with specific plans. The TCG develops and promotes provision for the protection of

computer resources from threats by malicious entities without infringing on the rights of end users. The Trusted Computing Platform provides cloud computing a sheltered base for achieve trusted computing [2]. A Trusted Platform Module (TPM) is a secure portal to potentially infinite amounts of protected storage, even though the time to store and retrieve particular information could ultimately become large.

2. THEORETICAL BACKGROUND
2.1 Cloud Computing
Cloud computing provides computation, software, data access, and storage services that do not necessitate end-user knowledge of the physical location and constitution of the system that delivers the services. The applications of cloud computing are practically unlimited. Through the right middleware, a cloud computing system could execute all the programs a ordinary computer could run. Everything from generic word processing software to personalized computer programs designed for a specific company could work on a cloud computing system [2].In a world that sees new technological trends blossom and fade on almost a daily basis, one new trend promises more prolonged existence. This trend is called cloud computing, and it will modify the way you use your computer and the internet. Cloud computing portends a major change in how we store information and run applications. Instead programs and data on an individual's desktop computer, everything is hosted in the "cloud"-an unformulated assemblage of computers and servers accessed via internet [2].

Cloud computing lets you access all your applications and document from anywhere in the world, baggage the confines of the desktop and making it easier for group members in different locations to collaborate. Cloud computing though it appears as network computing. With network computing applications or documents are hosted on a single company's server and accessed over the company's network. Cloud Computing starts getting different here. It encompasses various companies, various servers, and various networks. Cloud services and storage are accessible from anywhere in the world over the Internet connection. Cloud computing is also not an outsourcing process, where a company farms out (subcontracts) its computing services to an outside firm.

Cloud Computing is a growing method of Global Computing. Here the user can connect to the internet and start using all the required resources without a client side application installed on user's system. This eliminates the Physical storage mechanism on the client machine. Cloud computing differs from the classic client-server model by provide applications from a server that are executed and managed by a client's web browser. Centralization gives cloud service providers complete control over the version of the browser-based applications provided to clients, so no need for version upgrades or license management on individual client computing devices. Traditional business applications have always been very complex and expensive.

2.2 Trusted Computing
The Trusted Computing Group (TCG) proposed a set of hardware and software technologies to enable the construction of trusted platforms. The Trusted Computing Platform (TCP) will be used in authentication, confidentiality and integrity in cloud computing environment [14]. Trusted computing Platform is a computing platform that has a trusted component, most likely in the form of built-in hardware, which it uses to create a base of trust for software processes [4]. The Trusted Computing Group proposed a set of hardware and software technologies to enable the construction of trusted platforms. The Trusted Computing Platform will be used in authentication, confidentiality and integrity in cloud computing environment TC is controversial because it is technically possible not just to secure the hardware for its owner, but also to secure against its owner [18].

In recent years, increased confidence on computer security and the unfortunate fact of lack of it, particularly in the open-architecture computing platforms, have motivated many efforts made by the computing industry. In 1999, HP, IBM, Compaq, Intel, and Microsoft announced the formation of the Trusted Computing Platform Alliance (TCPA) that focused on building confidence and trust of computing platform in e-business transactions [15]. In 2003, the Trusted Computing Group was formed and has adopted the specifications developed by TCPA. Because one of the biggest issues facing computer technology today is data security, and the problem has gotten worse because users are working with sensitive information more often, while the number of threats is growing and hackers are developing new types of attacks, many technology researchers advocate development of trusted computing systems that integrate data security mechanism into their core operations, rather than implementing it by using add-on applications It is safer remote access through a combination of machine and user authentication and protects against data leakage by confirmation of platform integrity prior to decryption[18].

The Trusted Platform Module is an international standard, hardware security component built into many computers and computer-based goods. The TPM includes capabilities such as machine authentication, hardware encryption, secure key storage, and attestation. Encryption and signing are well known techniques, but the TPM makes them stronger by storing keys in protected hardware storage space. Machine authentication is a core principle that allows clouds to authenticate to a known machine to provide this machine and user a higher level of service as the machine is known and authenticated.

2.3 Trusted Computing Security Services
Trusted Computing Platform operates through a combination of software and hardware.TCP provides following security services,

Authenticated Boot

An authenticated boot service used to monitors what operating system software is booted on the computer and also tell which operating system is running. Each site in the cloud computing system will record the visitor's information. So by using the TCP mechanism in cloud computing, the trace of participants can be known by the cloud computing trace mechanism.

Encryption

Encryption is a process of translating the cipher text into plain text. This function lets data be encrypted in such a way that it can be decrypted only by a certain machine, and only if that machine is in a certain configuration. The encryption is another major mechanism in our design. This

service is built by a combination of hardware and software application.

Authentication

Authentication is the act of confirming the truth of an attribute of a datum or entity. Authentication provides the access permission to only the authorized users and restricts the unauthorized users.

Confidentiality

The information belongs to different owners in the cloud computing resources should be open to the trusted objects. Unauthorized people or other entities should be forbidden from that information.

Integrity

In integrity, cannot modify the originality of the information so integrity is regarded as the honesty and truthfulness or precision of one's actions. Integrity can be regarded as the opposite of duplicity, in that it regards internal consistency as a good feature, and suggests that parties holding apparently contradictory values should account for the inconsistency or alter their beliefs.

2.4 Trusted Components

Trusted computing consist of the following components,

Trusted Platform Support Services:

Trusted Platform Support Services is middleware that act as an intermediate between the TCP and the users.

Trusted Platform Module:

Trusted Platform Module is a security device that Can Store the cryptographic keys.

Core Root of Trust for Measurement:

It is software that can be used to identify the trusted root.

2.5 Need of Trusted Computing

With the ever increasing threat to identities and sensitive information, effective solutions can no longer be based on software only solutions, but on hardware which Trusted Platforms contain[17].Top problems and threats that a Trusted Platform can address:

- Identity theft and impersonation through unprotected passwords and sensitive information.
- Unauthorized network access, such as to a corporate network, a wireless network, or a VPN
- Regulatory compliance issues for strong authentication and data protection.
- Unauthorized access to unprotected files, documents, or email on client PCs or servers.

3. ARCHITECTURE

The architecture was designed to encompass a wide variety of tools and technologies. It provides strong user authentication, blocks the access of unsafe endpoints and coordinates security devices across the enterprise. The Trusted Computing technology is used to improve the security of cloud computing system.
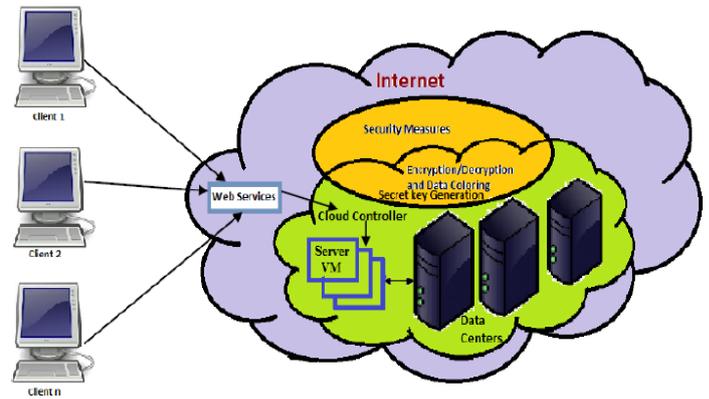
Fig 1. Trusted Cloud Computing

Trusted Computing was developed and promoted by the Trusted Computing Group. The word is taken from the field of trusted system*s* and has a specific meaning. We can generate the "master secret key" for each machine, and it uses the master secret to generate a unique sub-key for every possible configuration of that machine With Trusted Computing, the computer will constantly behave in expected ways and those behaviours will be enforced by hardware and software. This mechanism is a good choice because the web service technology has been well established in the network-computing environment. We can use the data colouring on varying security levels based on the variable function and apply the method to protect software, documents, images, video and relational databases. The details involved in the colour-matching process, which aims to combine a colured data object with its vendor, whose user identification is also colored with the same expected value depends on the data content, whereas *entropy* and *hyper entropy* add randomness or uncertainty. The colour-matching process assures that colours applied to the user identification match the data colours. This can initiate various trust-management events, including authentication, confidentiality, integrity and Virtual storage supports colour generation, embedding, extraction. Combining secure data storage and data colouring, we can prevent data objects from being damaged, stolen, altered and deleted.

4. TRUSTED COMPUTING TECHNOLOGY
4.1 Trusted Platform Model

TPM can implement security policies on hierarchies of secret keys to protect them from software attacks by some remote attacker. The Trusted Computing Platform Alliance (TCPA) has published documents that specify how a Trusted Platform must be constructed. Within each Trusted Platform is a Trusted (Platform) Subsystem, which contains a Trusted Platform Module (TPM), a Core Root of Trust for Measurement (CRTM), and support software.

The TPM is a hardware chip that's separate from the main platform CPU(s). The CRTM is the first software to run during the boot process and is preferably physically located within the TPM, although this isn't essential. The Trusted platform Support Service (TSS) performs a variety of functions, such as those necessary for communication with the relax of the platform and with other platforms [15].The TSS functions don't require to be trustworthy, but are however required if the platform is to be trusted. In addition to the Trusted Subsystem in the substantial Trusted

Platform, Certification Authorities (CAs) are centrally involved in the manufacture and usage of Trusted Platforms in order to guarantee that the TP is genuine [15]. Readers with a background in information security know that a Trusted Computing Base (TCB) is approximately the set of functions that provide the security properties of a platform.

The TCB in a Trusted Platform is the combination of the Trusted Subsystem (mainly dealing with secrets) and additional functions. As such, the Trusted Subsystem is a subset of the functions of the Trusted Computing Base of conventional sheltered computers, which would normally include both dealing with secrets and using secrets. Crucially, however, the Trusted Subsystem contains some functions not found in a conservative TCB. Conventional secure computers provide formal proof that a TCB in certain states actually can be trusted [15].

4.2 Trace of the User's Behaviours

The users have chock-full information about their identity, the cloud computing system can use some mechanism to trace the users and get their source. Since in the TCP the user's identity is proved by user's special key and this mechanism is included in the hardware, such as the BIOS and TPM. It is very hard to the user to make unreliable for their identity information. Previous to the distributed machine cooperates to do a little, they should attest their neighbouring information to the remote location. When the user login the cloud computing system and his identity information should be recorded and verified at foremost. The cloud computing security can be provided as security services and resources. Security messages and secured messages can be transported, unstated, and manipulated by model Web services tools and software. This mechanism is a good choice because the web service technology has been well established in the network-computing environment. Even the mechanism for the cloud computing security has many merits now, but there are still a number of disadvantages.

Each site in the cloud computing system will record the visitor's information. So if the TCP mechanism is integrated into the cloud computing, the trace of the participants, including the users and other resources, can be knew by the cloud computing trace mechanism. Then if the participants do some malicious behaviour, they will be tracked and be punished. In order to achieve the trusted computing in the cloud computing system, we should have the mechanism to know not only what the participants can do, but also what the participant have done. So the monitoring function should be integrated into the cloud computing system to supervise the participants' behaviour. In reference monitors have been used in the operation system for more than several decades.

4.3 Cloud Computing based on TCP

The trusted computing technology can provide a way that can help to set up a security cloud computing background. The model of trusted computing is initially designed to provide the privacy and trust in the personal platform and the trusted computing platform is the base of the trusted computing. Because the internet computing or network computing has been the main computing from the ending of the last century, the model of trusted computing is

being developed to the network computing, especially the distributed systems environment.

The cloud computing is a promising distributed system model and will act as an important role in the e-business or research environments. As web service technology have developed quickly and have been used broadly, cloud computing system could evolve to cloud computing service, which integrates the cloud computing with web service technology. So we could extend the trusted computing mechanism to cloud computing service systems by integrating the TCP into cloud computing system. Trusted computing platform provide the basis for trusted transactions to occur, and trusted computing technologies must allow stakeholders to express policies and have those policies negotiated and enforced in any execution environment.

4.4 Trusted Hardware

Trusted Computing technology as it exists nowadays is distinct by the specifications of the Trusted Computing Group. Hardware component, the Trusted Platform Module, is integrated into commonly available general-purpose hardware, with millions of platforms shipped so far away. Like a smart card, a Trusted Platform Module features cryptographic primitives, but it can be physically bound to its host device. Trusted hardware contains a tamper-resistant integrated circuit implementation public key cryptography, key generation, secure hashing, and random-number generation. To use these components, the Trusted Platform Module (TPM) can enforce security policies on hierarchies of secret keys to protect them from software attacks by any remote attacker.

Trusted Platform Module can be used to perform cryptographic signatures on user-provided data using hardware-protected private keys. However, due to limited TPM memory, keys have to be swapped out of the TPM when not in use. To protect these keys, a parent storage key specified on key creation is used to wrap the private part of the child key when it is exported from the TPM. At the top of the key hierarchy is the storage root key created when taking ownership of the TPM. Keys are assigned a user-supplied secret, which is used in several authentication protocols, and optionally a system state that has to be provided when using the key for cryptographic operations [12].

4.5 Authentication of cloud computing environment with Trusted Computing Platform

Data protection is a more than just a subject of maintenance in the wrong people out of places they shouldn't be and not having valuable records disappear. Data protection is a driven by a host of new legal requirements that protect the customer privacy. It is Critical to data protection will be the safe linking of host CPU and hard drives. Different entities can appeal to join the cloud computing environment. The initial step is to verify their identities to the cloud computing system administration.

Because cloud computing should involve a large amount of entities, such as users and resources from different sources, the authentication is important and complicated. Considering these, we use the TCP to aid to process the authentication in cloud computing. The TCP is based on the TPM. The TPM is a logic independent hardware [18]. It can resist the attack from software, and even the hardware attack. The TPM contain a private master key which can provide

protect for other information store in cloud computing system. Because the hardware certificate can store in TPM, it is hard to attack it. So TPM can provide the trust root for users.

The cloud computing service should present which role it will give the permission, when the cloud computing service notifies itself to the cloud -computing environment. So the user will able to know whether he could make access to that cloud computing service before his action. The encryption is another major mechanism in our design. This function lets data be encrypted in such a way that it can be decrypted only by a certain machine, and only if that machine is in a certain configuration. This service is built by a combination of hardware and software application. The hardware maintains a "master secret key" for each machine, and it uses the master secret to generate a unique sub-key for every possible configuration of that machine. As a result, data encrypted for a particular configuration cannot be decrypted when the machine is in a different configuration.

When one machine wants to join the cloud computing, it will show its certificate and generate session key with other co-operators buy using the unique sub-key. If the configuration in the local machine is changed, the session-key will also be not useful. So in the distributed environment, we can use this function to transmit data to remote machine and this data can be decrypted when the remote machine has certain configuration.

4.6 Trusted platform Support Service

TSS components are the major parts of the TCP enabled cloud computing. It provides fundamental resources to support the TPM. In our design, TSS should be a bridge between the up-application and the low-hardware. Trusted platform Support Service (TSS) includes two layers, the TSS service provider (TSP) and TSS core services (TCS). The applications call the function of TSP. TSP provides some basic security function modules. These basic modules send calls to TCS. Then TSS converts these calls to according TPM instructions. Since TPM is hardware, the TCG Device Driver Library (TDDL) is necessary. TDDL convert the calls from TCS to the TPM orders.

After the TPM process the order, it will return the results up forward. Each layer gets results from low layer and coverts them to responding results that the up layer needs.

The main issue with the "Cloud" is linked to the responsiveness of information. In a cloud, each of us is completely right to be concerned about the confidentiality and the availability of the information. Tomorrow's world will be based on information. But the future is becoming gradually more doubtful. Information is a critical resource that requires severe controls and protection.

4.7 Trusted Computing Benefits

Trusted Computing technology creates a safer environment in cloud computing. It provides Safer Remote Access through a Combination of mechanism and User Authentication. Trusted computing Protects against data leakage by confirmation of platform integrity prior to encryption and decryption. The Hardware Protection for Encryption and Authentication Key is used by Data (Files) and Communications (Email, Network Access). The Hardware Protection for individually Identifiable

Information such as User Ids and Passwords. Lowest Cost Hardware Security Solution: No Token to Distribute or Lose, No Peripheral to Buy or Plug In, No Limit to Number of Keys, Files or IDs Protected.

- Trusted Computing Protect Business Critical Data and Systems.
- Secure Authentication and Strong Protection of User IDs.
- Establish Strong Machine Identity and Integrity.
- Ensure Regulatory Compliance with Hardware-Based Security.
- Trusted Computing Reduce the Total Cost of Ownership through "Built In" Protection.

5. CONCLUSION AND FUTURE WORK

This paper analyzed and finds the role of trusted computing platform in cloud computing. Trusted Computing Platform is used as the hardware foundation for the cloud computing system. Trusted Computing Platform provides cloud computing system with some imperative security functions, which include authentication, confidentiality, integrity, communication security and data protection.

The advantages of our planned approach are extending the trusted computing technology to accomplish its requirements for the cloud computing and then fulfil the trusted cloud computing. To integrate these hardware modules with cloud computing system is a difficult work and need more unfathomable study. We develop a model system of trusted cloud computing, which is based on the trusted computing platform. It can provide stretchy security services for users. The Trusted Computing Platform provides cloud computing a sheltered base for achieve trusted computing. We will make the actual design more practical and operational in the imminent. In future, we would also like to study over the impact of more security in this proposed method.

REFERENCES

[1] Balachandra Reddy Kandukuri, Ramacrishna PaturiV, Atanu Rakshi, "Cloud Security Issues",IEEE International Conference on Services Computing, pages(s):517-520, 2009.

[2] CloudComputing:http://en.wikipedia.org/wiki/Cloud_computing , Accessed: 28/07/2011.

[3] Cloud Computing, http://www.techno-pulse.com/ Cloud Computing for Beginners, Accessed: 28/07/2011.

[4] Cloud Security Alliance: Security Guidance Critical Areas of Focus in Cloud Computing, http://www.cloudsecurityalliance.org/guidance/csaguide.pdf. April 2009.

[5] Dr.Rao Mikkilineni, Vijay Sarathy, "Cloud Computing and the Lessons from the Past", the 18th IEEE international Workshops on Enabling Technologies: Infrasturctures for Colloaborative Enterises, on page(s):57-62, 2009.

[6] Frank E. Gillett, "Future View: The new technology ecosystems of cloud, cloud services and cloud computing" Forrester Report, August 2008.

[7] Glen Bruce, Rob Dempsey, "Security in Distributed Computing", Published by Prentice Hall, Copyright Hewlett-Packard Company, 1997.

[8] ISO/IEC. Information technology-Open Systems Interconnection- Evaluation criteria for information tech-nology, Standard ISO/IEC 15408.1999.

[9] Jason Reid Juan M. González Nieto Ed Dawson, "Privacy and Trusted Computing", Proceedings of the 14th International Workshop on Database and Expert Systems Applications, IEEE, 2003.

[10] Martín Abadi, "Logic in Access Control", Proceedings of the18[th] Annual IEEE Symposium on Logic in Computer Science (LICS'03), 2003.

[11] Peter Wayner, "Cloud versus cloud – A guided tour of Amazon, Google, AppNexus and GoGrid", InfoWorld, July 21, 2008.

[12] Ronald Toegl, Thomas Winkler, Mohammad Nauman, Theodore Hong, "Towards Platform-Independent Trusted Computing",2009.

[13] Tal Garfinkel, Mendel Rosenblum, and Dan Boneh, "Flexible OS Support and Applications for Trusted Computing", the 9th Workshop on Hot Topics in Operating Systems (HotOS IX), USENIX, 2003.

[14] Trusted Computing Group (TCG), "TCG Specification Architecture Overview Specification Revision 1.2", April 28, 2004.

[15] Trusted computing group: http://www.trustedcomputinggroup.org. Accessed: 28/07/2011.

[16] Trusted computing Technology : http://en.wikipedia.org/wiki/Trusted_Computing. Accessed: 28/07/2011.

[17] Trusted computing : http://www.wave.com. Accessed: 30/07/2011.

[18] Zhidong Shen, Qiang Tong, "The Security of Cloud Computing System enabled by Trusted Computing Technology", Proceedings of the 2nd International Conference on Signal Processing Systems (ICSPS), 2010.

ABOUT AUTHORS

Mr. P.Senthil is currently pursing M.E, CSE in Coimbatore Institute of Engineering and Technology (C.I.E.T), Coimbatore and received his B.Tech degree in Information Technology at Dhanalakshmi Srinivasan Engineering College, Perambalur. His Research areas are Networking and Cloud Computing.

Mr.N.Boopal is working as an Assistant Professor in C.I.E.T, Coimbatore. He received his M.E degree in Computer Science and Engineering at Anna university of Technology, Coimbatore and received B.E CSE in Coimbatore Institute of Engineering and Technology (C.I.E.T), Coimbatore. He has a four and half years of teaching experience. His Research areas are Mobile Cloud Computing and Software Engineering.

Ms R.Vanathi is currently pursing M.E, CSE in Coimbatore Institute of Engineering and Technology (C.I.E.T), Coimbatore and received her B.E Degree in Information Technology at Karunya University, Coimbatore. Her Research areas are Networking and Cloud Computing.

Protection Of Power System By Optimal Co-ordination of Directional Overcurrent Relays Using Genetic Algorithm

Dharmendra Kumar Singh[*], Dr. S. Gupta[**]

*(M.Tech, Student, Department Of Electrical Engineering, NIT Raipur, India)
** (Department Of Electrical Engineering, NIT Raipur, India)

ABSTRACT

In this paper, the optimization of coordination of directional overcurrent relays in an interconnected power system is presented. The objective of protective relay coordination is to achieve selectivity without sacrificing sensitivity and quick fault clearance time. A methodology is adopted for the consideration of backup relays in the optimal coordination of directional overcurrent relays. The calculation of the Time Dial Setting (*TDS*) and the pickup current (*Ip*) setting of the relays is the core of the coordination study. The objective function to be minimized is defined as the sum of the time dial settings of all the relays. The inequality constraints guarantee the coordination margin for each primary/backup relay pair having a fault very close to the primary relay. Using this formulation, the size of the optimization problem is significantly reduced. Genetic Algorithm is the algorithm being applied to minimize the operating times of the relays. Both Linear and Non Linear Equations are framed for the test bus system used and optimized using the Genetic Algorithm in this paper.

Keywords - Directional Overcurrent Relay, Genetic Algorithm, Pickup Current, Time Dial Setting, Operating Time of Relays.

1. INTRODUCTION

IN an interconnected power system, abnormal conditions (faults, overload, overvoltage, etc.) can frequently occur. Due to this, interruption of the supply and damage of equipments connected to the power system may occur. During these situations, the faulted components must be readily identified and isolated in order to guarantee the energy supply to the largest number of consumers possible and to maintain the system stability. Therefore a reliable protective system is required. To ensure reliability, a backup protective system should exist in-case the main protective system fails (relay fault or breaker fault). This backup protection should act as a backup either in the same station or in the neighboring lines with time delay according to the selectivity requirement. Directional over current relay(DOC) are commonly used for power system protection. Optimization method are used for coordination of these relays. In these methods, at first, the coordination constraints for each main and backup relay are determined [1]-[2]. The other optimization methods that have been used for the above problems are simplex method [3] two phase simplex method [4] and dual simplex [5].

1.1 Problems Associated With Above Techniques

A protective relay should trip for a fault in its zone and should not, for a fault outside its zone, except to backup a failed relay or circuit breaker. The problem of coordinating protective relays in electric power systems consists of selecting their suitable settings such that their fundamental protective function is met under the requirements of sensitivity, selectivity, reliability and speed.

1.2 Justification for Using Genetic Algorithms

This paper describes a systematic Overcurrent (OC) protection grading method based on GENETIC ALGORITHM (GA). Genetic Algorithms are computerized search and optimization algorithms based on the mechanics of natural genetics and natural selection. Professor Holland of University of Michigan envisaged the concept of these algorithms in the mid 60's. Thereafter a number of students and other researchers have contributed to the development of this field. Genetic Algorithms are good at taking larger, potentially huge, search space and navigating them looking for optimal combinations of things and solutions which we might not find in a life time.

1.3 Solution Approach

In [5], for Linear Programming, the values of Time Dial Setting (TDS) have been found for given values of pickup currents (Ip) and in [4], for Non-linear Programming, the values of pickup currents (Ip) have

been found for given values of TDS using GA subject to the constraints and hence the operating times of the relays is minimized.

2. OPTIMAL COORDINATION PROBLEM

Directional Overcurrent Relay (DOCR) coordination problem is a parametric optimization problem, where different constraints have to be considered in solving the objective function [1]-[2]. Here the objective function to be minimized is the sum of the operating times of the relays connected to the system, subject to the following constraints.

2.1 Relay Characterstics

A typical Inverse Time Directional Over current relay consists of two elements, an instantaneous unit and a time overcurrent unit. The overcurrent unit has two values to be set: Pick up current value (I_p) and the Time Dial setting (TDS)/TMS. The pickup current value (I_p) is the minimum current value for which the relay operates[4]. The Time Dial Setting defines the operation time (T) of the device for each current value, and is normally given as a curve T vs M, where M is the ratio of relay fault current I, to the pickup current value:

$$M = I/I_p \qquad (1)$$

In general, overcurrent relays respond to a characteristic function of the type:

$$T = f(TDS, I_p, I) \qquad (2)$$

This Function can be approximated as:

$$T = \left\{ \dfrac{K_1 * TDS}{\dfrac{I * K_2}{CT\ Ratio * I_p + K_3}} \right\} \qquad (3)$$

Where K_1, K_2 and K_3 are constants that depend upon the specific device being simulated.

The following two cases have been considered in this paper for obtaining the objective function and minimizing it.

Case I

The Linear equation is formulated in terms of TDS by taking Ip to be constant at a particular value in equation (3)

Case II

The Non-Linear equation is formulated in terms of Ip by taking TDS to be constant at a particular value in equation (3).

2.2 Relay Settings

The calculation of the two settings, TDS and I_p, is the essence of the directional Overcurrent relay coordination study. It is very important to mention that in general, directional overcurrent relays allow for continuous time dial settings but discrete (rather than continuous) pickup current settings.

Therefore this constraint can be formulated as:

$$TDSi_{min} \leq TDSi \leq TDSi_{max}$$

$$I_{Pimin} \leq I_{Pi} \leq I_{Pimax}$$

2.3 Coordination Problem

In any power system, a primary protection has its own backup one for guaranteeing a dependable power system. The two protective systems (primary and back-up) should be coordinated together. Coordination time interval (CTI) is the criteria to be considered for coordination. It's a predefined coordination time interval and it depends on the type of relays. For electromagnetic relays, CTI is of the order of 0.3 to 0.4 s, while for a microprocessor based relay, it is of the order of 0.1 to 0.2 s.

To ensure reliability of the protective system, the back-up scheme shouldn't come into action unless the primary (main) fails to take the appropriate action. Only when CTI is exceeded, backup relay should come into action.

This case is expressed as:

$$T_{backup} - T_{primary} \geq CTI$$

Where, T_{backup} is the operating time of the backup relay

$T_{primary}$ is the operating time of the primary relay

After considering all these criteria, this problem can be formulated mathematically as:

$$\min \sum_{i=1}^{n} T_i$$

Where, n represents the number of relays.

2.4 Optimization Of TDS & I_P

The objective function is found and various coordination time constraints are formulated as proposed earlier. Coordination time constraints are formulated from the relay pair tabulation.

The optimization of the functions is carried out by the GA toolbox by entering data in the appropriate fields.

The objective function and constraints for the three bus and six bus systems are formulated. The proposed methodology has been applied to a 3-bus test system. The coordination time interval of 0.2 seconds is used. The TDS values ranges from 0.1 to 1.1 and I_p from 0.1 to 1.1. The optimum TDS and I_p values are determined using Genetic Algorithm.

For Six bus systems, the P/B pairs and the fault currents for the systems are determined as per the proposed methodology. The TDS are assumed to vary between a minimum value of 0.5 to 1.1 and also I_p from 0.5 to 1.1. A CTI of 0.2 is taken, while the transient changes in the network topology are not considered. The optimum TDS and I_p values are determined using Genetic Algorithm.

3. GENETIC ALGORITHM

Genetic algorithm (GA) is a search technique used in computing, to find exact or approximate solutions to optimization and search problems. Genetic algorithms are categorized as global search heuristics [6]-[7]. Genetic algorithms are a particular class of evolutionary algorithms (EA) that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover[8]-[9].

3.1   Terminologies Related To Genetic Algorithm

- Fitness Function--A fitness function is a particular type of objective function that prescribes the optimality of a solution (that is, a chromosome) in a genetic algorithm sothat particular chromosome may be ranked against all the other chromosomes.
- Chromosome-- In genetic algorithms, a chromosome is a set of parameters which define a proposed solution to the problem that the genetic algorithm is trying to solve. The chromosome is often represented as a simple string, although a wide variety of other data structures are also used.
- Selection--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 are typically more likely to be selected.
- Reproduction--The next step is to generate a second generation population of solutions from those selected through genetic operator crossover.

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.

- Crossover--In genetic algorithms, crossover is genetic operator used to vary the programming of a chromosome or chromosomes from one generation to the next. It is analogous to reproduction and biological crossover, upon which genetic algorithms are based.
- Mutation--In genetic algorithms of computing, mutation is a genetic operator used to maintain genetic diversity from one generation of a population of algorithm chromosomes to the next. It is analogous to biological mutation.

3.2   Outline Of The Genetic Algorithm

It is very clear from flow chart given in Fig. 1.
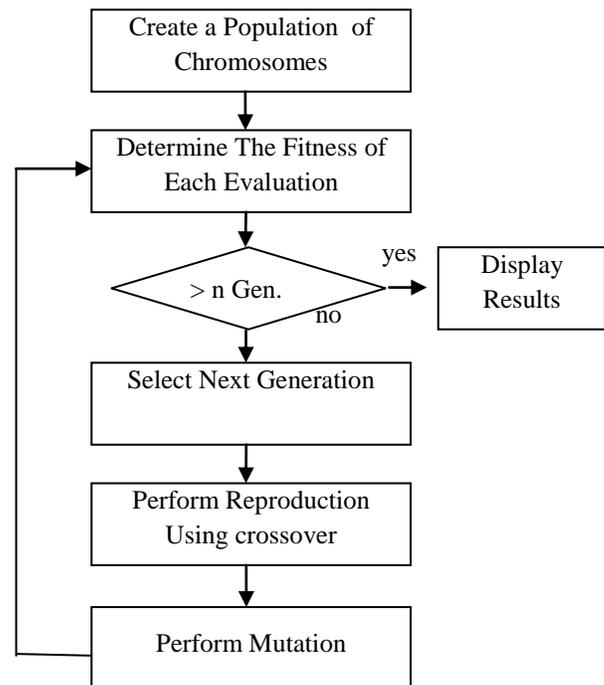
Fig. 1. Flow chart

4. PROPOSED ALGORITHM

Step 1: Generate P/B relay pairs. The knowledge of primary/backup relay pairs is essential in the formulation of the coordination constraints.

Step 2: Load flow analysis is done using Newton-Raphson method to determine line currents. This analysis can be done using any simulation softwares and here we have used ETAP simulation software .

Step 3: Short circuit analysis is done using the same simulation software to find fault currents.

Step 4: Pickup current is calculated based on the load current. Here it will be set at 1.5 times the maximum load current, for phase protection.

Step 5: Minimization of objective function is carried out and optimum values of TDS are determined using linear programming technique in MATLAB.

Step 6: The values of TDS and minimization of objective function is further optimized using Genetic Algorithm (GA) using MATLAB Toolbox.

5. SIMULATION AND RESULTS

The performance of the proposed method is evaluated by, a 3-bus overcurrent relay network shown in Fig. 2. All relays have inverse characteristic and the aim is to find an optimal setting for network relays in order to minimize the final operating time and in such a way that all constraints are satisfied. The coordination time interval (CTI) is assumed to be 0.2 seconds. In order to solve the relay coordination problem first, it needs to determine the primary and backup relay pairs.

For determining the primary and backup relay pairs, first select a relay then determine the relays which install in the far bus of the selected relay, then omitting the relay which is in the same line with selected relay. Now, it can say that the selected relays are the backup relays for the primary relay. Table I shows the primary/backup relays pairs and faults currents. The ratios of the current transformers (CTs) are indicated in Table II.

Fig. 3. shows the convergence of the proposed method and the algorithm converge to the global solution. The result of the proposed method are shown in Table III and Table IV for linear and nonlinear function respectively.
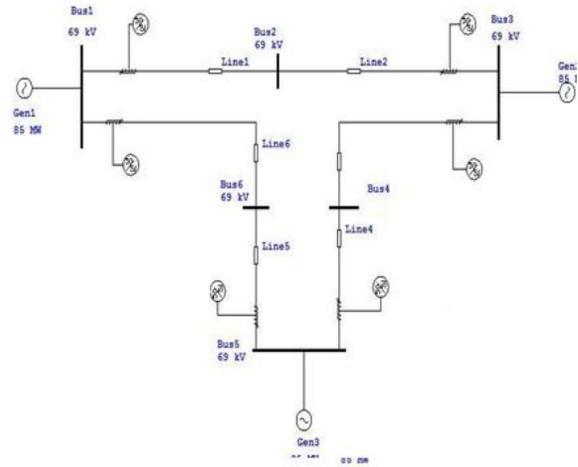
Fig. 2. Three bus system

Table I
Primary/Backup Relay Pairs And Fault Currents

Backup Relay	Fault current (KA)	Primary Relay	Fault current (KA)
5	1.457	1	2.673
4	0.327	2	1.666
1	0.543	3	1.666
6	0.765	4	2.673
3	1.005	5	1.325
2	0.670	6	2.469

Table II
CT Ratio

Relay no.	CT Ratio	Relay no.	CT Ratio
1	400/5	4	400/5
2	200/5	5	300/5
3	200/5	6	200/5
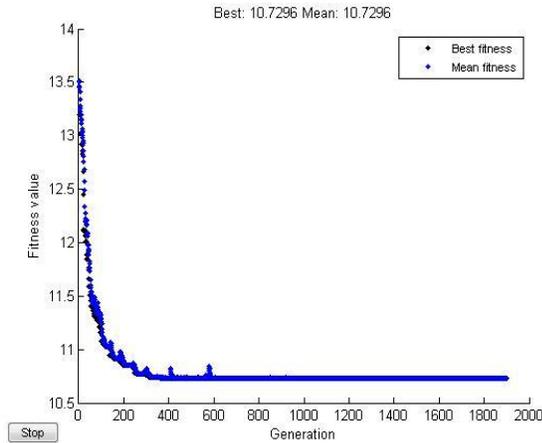
Fig. 3 Convergence of the proposed method

Table III
Result Using GA (For Linear Function)

Time dial Setting	Value Obtained
TDS$_1$	0.10001
TDS$_2$	0.18023
TDS$_3$	0.11908
TDS$_4$	0.12004
TDS$_5$	0.15236
TDS$_6$	0.1192
Fitness Function Value= 1.69083 sec	

Table IV
Result Using GA(For Nonlinear Function)

Pick up Current	Obtained Values
I_{p1}	0.2831
I_{p2}	0.60123
I_{p3}	0.10091
I_{p4}	0.42051
I_{p5}	1.0808
I_{p6}	0.5355
Fitness Function Value= 1.22844 sec	

6. CONCLUSION

In this paper, an optimization methodology is presented to solve the problem of coordinating directional overcurrent relays in an interconnected power system. The operating time of the relays was determined using GENETIC ALGORITHM for 3 bus system for both Linear Objective Functions and for Non Linear Objective Functions. The value of Time Dial setting (TDS) and Pick up current value (I_p) are found for linear and nonlinear function respectively in such a way that all constraints are satisfied. This method increases coordination and the operation speed of relays. Finding the absolute optimal point, the ability to be apply on large networks, the ability to consider both linear and non-linear characteristics of relays are some of advantages of the proposed method.

The dependency of GA solution on initial condition is weaker and it requires more computing time. These are the some limitations of this method.

REFERENCES

Books:

[1]   P.M. Anderson, *Power system protection*. Newyork : McGraw-Hill,1999.

Proceedings Papers:

[2]   L.Perez, A.J.Urdaneta, "Optimal Computation of Directional Overcurrent Relays Considering Definite Time Backup Relaying ", IEEE Transaction on Power Delivery, Vol. 4, No.4, October 1999, PP. 385-388

[3]   L.Perez, A.J.Urdaneta, "Optimal Computation of Distance Relays Second Zone Timing in a Mixed Protection Scheme with Directional Relays", IEEE Transaction on Power Delivery, Vol. 16, No.3, July 2001, PP. 1276-1284.

[4]   AI. Urdaneta, R. Nadira and L.G. Perez Jimenez, "Optimal Coordination of Directional Overcurrent Relays in Interconnected Power Systems," IEEE Trans. Power Del, Vol 3, No. 3, PP 903-91 I, July.1988.

[5]   B. Chattopadhyay, M. S. Sachdev and T.S. Sidhu, "An On-Line Relay Coordination Algorithm for Adaptive Protection Using Linear Programming Technique," IEEE Trans. Power Del, Vol 11, No. 1, PP 165-173, Jan. 1996.

[6]   Cheng-Hung Lee and Chao-Rong Chen - Using Genetic Algorithm for overcurrent Relay

Coordination in Industrial Power System - The 14th National Conference on Intelligent Systems Applications to Power Systems, ISAP 2007 National Taipei University of Technology, Taiwan.

[7] C. W. So, K. K. Li, K. T. Lai and K. Y. Fung, "Application of Genetic Algorithm to Overcurrent Relay Grading Coordination," APSCO. 97, Vol 1, PP. 283-287 ,Nov. 1997.

Journal Papers:

[8] Farzad Razavi, Hossein Askarian Abyaneh - A new comprehensive Genetic Algorithm method for Optimal overcurrent relays coordination – International Journal on Electric Power Systems Research, 78(2008) Pg: 713-720.

Theses:

[9] David E. Goldberg - Genetic Algorithms in Search, optimization and Machine Learning - Addison-Wesley, Reading MA,

An Overview of Disarray in Nonlinear Active Suspension System under Random Road Excitation with Time Delay

N. R. Kumbhar[1], S. H. Sawant[2], S.T. Satpute[3], Dr. J.A. Tamboli[4]

[*]PG Student, Dept. of Automobile Engineering, Rajarambapu Institute of Technology, Sakharale. Shivaji University, Kolhapur, India

[**]Dept. Of Mechanical Engineering, Dr. J. J. Magdum College of Engineering, Jaysingpur. Shivaji University, Kolhapur, India

[***] Dept. of Automobile Engineering, Rajarambapu Institute of Technology, Sakharale. Shivaji University, Kolhapur, India

[****] Principal, Annasaheb Dange College of Engineering, Ashta,Shivaji University, Kolhapur, India

ABSTRACT:

Since the vehicle dynamics is concerned with controllability and stability of vehicle, it is important in design of a ground vehicle. The modeling of the vehicle with the analysis of the dynamic response of the mathematical model have been examined in a large number of investigations. In this paper overview of various works are done. This paper tries to give an idea about the previous researches & their finding about study of nonlinearity of active suspension system parameters by considering half car model. Some times help of quarter car model is taken for more detail understanding of system behavior.

Keywords – **Active suspension, Chaos and Bifurcation, Random road excitation, Time delay.**

I. INTRODUCTION

A passive suspension system (PSS) is designed to preserve two desired aims which are vehicle handling and passenger comfort. A design problem is to provide a trade off between the both aims which are opposite to one another. The PSS cannot adapt itself in the face of wide changes in road conditions. However, this can be accomplished by controlling vertical acceleration of a vehicle using an active suspension system (ASS). It comprises a force generating actuator placed between sprung and unsprung mass.

During the design of a suspension system, a number of conflicting requirements has to be met. The suspension setup has to ensure a comfortable ride and good cornering characteristics at the same time. Also, optimal contact between wheels and road surface is needed in various driving conditions in order to maximize safety. Instead of a passive suspension, in

most of today's cars, an active suspension can be used in order to better resolve the trade-off between these conflicts. However, this is generally accompanied by considerable energy consumption. In this paper literature on an active suspension is discussed. Since the disturbance from the road may induce uncomfortable shake and noise in the vehicle body, it is important to study the vibrations of the vehicle. Many studies have been carried out on the dynamic response and the vibration control with linear mechanical model. However, an automobile is a nonlinear system in practice because it consists of suspensions, tires and other components that have nonlinear properties. Therefore, the chaotic response may appear as the vehicle moves over a bumpy road.

By considering all above facts, this paper tries to cover literature which deals with an active suspension system with half car model by considering suspension nonlinearities and time delay parameter. Some times in this paper reference of quarter car model for above situation is considered to elaborate the concept in simple manner.

II. SLIDING MODE CONTROLLER FOR VEHICLE ACTIVE SUSPENSION SYSTEM

C Kim & P I Ro in year 1998 studied the *Sliding mode controller for vehicle active suspension system with nonlinearities* in this paper he investigated the control of an active suspension system using quarter car model. Due to nonlinearities in the real suspension system it is very difficult to achieve desired performance of suspension system. To ensure the robustness for wide rane of operating conditions a sliding mode controller was designed based on Sliding Mode theory and compared with existing nonlinear adaptive control scheme. With simulation results it was shown that both ride quality & handling

performance are improved using SM active suspension system in the preference of nonlinearities of suspension system and uncertainties of suspension parameters. The simulation results of SM controller scheme then compared with Self Turning Control (STC) scheme which was developed to deal with nonlinearities of suspension system. Simulation results shows that the both controller can improve the ride quality where as only SM active suspension system shows robust tracking performance even when suspension parameters changed suddenly.[1]

III. STOCHASTIC OPTIMAL PREVIEW CONTROL

In the year 2003 Javad Marzbarad, Goodarz Ahmadi, Hassan Zohoor, Yousef Hojjat made study on *Stochastic optimal preview control of vehicle suspension on random road*. The road roughness height is modeled as a filtered white noise stochastic process and a four-degree-of-freedom half-car model is used in the analysis. The suspension system is optimized by minimizing the performance index containing the mean-square values of body accelerations (including effects of heave and pitch), tire deflections and front and rear suspension rattle spaces. The effect of delay between front and rear wheels is included in the analysis. Responses of a vehicle suspension with active, active and time delay, and active and preview control systems to random road input are evaluated and the results are compared with those for the passive system. Road surface elevation information at distance ahead of the bumper is used as preview and the suspension space velocities are measured by sensors in a noisy environment. Stochastic optimal control theory is used and the states are estimated by an observer, similar to a Kalman filter. It is recognized that the actuator dynamics affects the active control system performance with or without the preview. In manuscript they have shown that the preview information improves the control system performance. One advantage of the preview control approach is that it can compensate for the time delays in the reactions of the system and the actuators; however, interactions of the actuation system with the suspension are not considered. For optimal priview control formulation they considered theorem that is given a system with state equation and preview time tp, that is with $w(\sigma)$, $\sigma\varepsilon[t+tp]$ known, the problem is to find a control law $u(t)=f(x(t), \; w(\sigma), \; \sigma\varepsilon[t+tp] \;)$ that minimizes the quadratic performance of index. [2]

IV. CHAOS AND BIFURCATIONS IN A NONLINEAR VEHICLE MODEL

Also Q. Zhu and M Ishitobi in 2003 had studied the *Chaos and bifurcations in a nonlinear vehicle model* for their study of chaotic responses and bifurcations of a four-degree-of-freedom vehicle model that is subjected to two sinusoidas disturbances with time delay are studied through numerical simulation. They found that the chaotic response may appear in the instable region of frequency-response diagram. The bifurcation diagram shows that the chaotic response could be sensitive to variation of damping of the suspension. To identify the chaotic motion of the system, dominant laypunov exponent is used. Although the mechanical model of the vehicle is only a simplified one and the parameters used do not agree closely with the practical data for an automobile, the results may still be useful in dynamic design of the ground vehicle.[3]

V. CONTROL OF RANDOM RESPONSE OF A HALF-CAR VEHICLE MODEL

In 2007 L.V.V. Gopala Rao, S. Narayanan *Preview control of random response of a half-car vehicle model traversing rough road* an active control of response of a four degree-of-freedom (dof) half-car model traversing at constant velocity a random road with look ahead preview is considered. The suspension spring is assumed to be hysteretic nonlinear and modeled by the Bouc–Wen model. The statistical linearization technique is used to derive an equivalent linear model. The response of the vehicle is optimized with respect to suspension stroke, road holding and control force. The RMS values of the suspension stroke, road holding and control forces are computed using the spectral decomposition method. The results for the equivalent linear model obtained by the spectral decomposition method are verified using Monte Carlo Simulation (MCS). And come to the result that the overall performance of the vehicle improves with preview control as compared to the performance without preview control. But the improvement with performance saturates beyond a preview distance. The equivalent linearization model and the equivalent step road input assumption are validated by MCS by generating the road input compatible with the power spectral density function and numerically integrating the equations of motion with the nonlinear Bouc–Wen model. [4]

VI. H∞ CONTROL OF ACTIVE SUSPENSION FOR A HALF CAR MODEL

In 2008 H Du and N Zahang presented work on *Constrained H∞ control of active suspension for a half-car model with a time delay in control*. The time delay for the control input is assumed to be

uncertain time invariant within a known constant bound.

Active half car model which is considered for study is shown in Fig1.
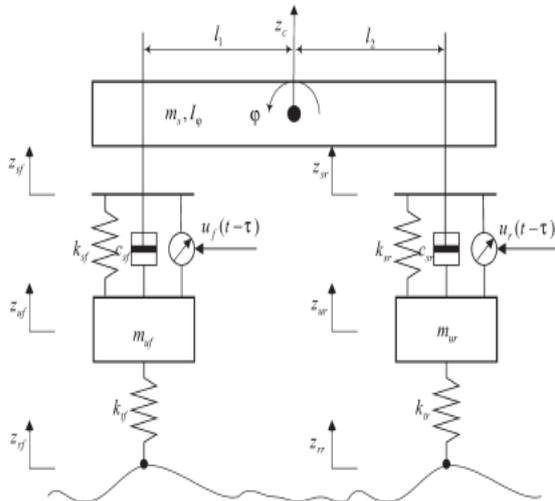
Fig 1: Active half car model

The ride comfort performance of the suspension system is optimized by using the H∞ norm to measure the body accelerations (including both the heaving and the pitching motions) , while the tire deflections and the suspension rattle spaces are constrained by their peak response values in time domain using the generalized H2 (GH_2) norm (energy-to-peak) performance.

Then, a constrained delay-dependent H∞ state feedback controller is designed to realize the ride comfort, road holding and stroke limitation performance to prescribed level in spite of the existence of a time delay in control input. And finds that the designed controller can keep the stability of the closed-loop system while improving the performance on ride comfort, keeping suspension strokes within given bounds, and ensuring firm contact of wheels to road even with the existence of a time delay in the control input to some extent.[5]

VII. IMPEDANCE CONTROL OF AN ACTIVE SUSPENSION SYSTEM

In the same year 2008 Mohammad Mehdi Fateh and Seyed Sina Alavi worked on the *Impedance control of an active suspension system.* They developed novel control system to control dynamic behavior of a vehicle subject to road disturbances. The novelty of their work was to apply the impedance control on an active vehicle suspension system operated by a hydraulic actuator. A relation between the passenger comfort and vehicle handling is derived using the impedance parameters. The impedance control law is simple, free of model and can be applied for a broad range

of road conditions including a flat road. Impedance control is achieved through two interior loops which are force control of the actuator by feedback linearization and fuzzy control loop to track a desired body displacement provided by the impedance rule. The system stability is analyzed. A quarter-car model of suspension system and a nonlinear model of hydraulic actuator are used to simulate the control system.The simulation results were presented to show the performance of control system by comparing the ASS and the PSS. Based on the simulation, they concluded that the impedance control of ASS was performed well preferred to PSS. In comparison with model based control laws such as optimal control law, the Ipedance Rule (IR) shows important advantages [6].

VIII. BIFURCATION, CHAOS AND CHAOTIC CONTROL OF VEHICLE SUSPENSION

Again in 2008 Yung sheng, Guang-qiang, Xian-jie Meng made study on *bifurcation, chaos and chaotic control of vehicle suspension with nonlinearities under road excitation* they obtained mathematical model for nonlinear suspension model from actual measured data of car by applying it to simulate the nonlinear spring force and damping force of vehicle suspension system. Model used for analysis is shown in Fig. 2
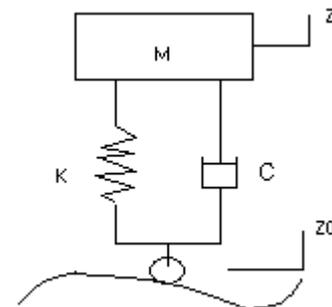
Fig2: Quarter Car Single DOF Model

A dynamic equation of quarter suspension system with s.d.f. is derived. The theory of nonlinear vibration is applied to study the nonlinear model and to reveal its nonlinear vibration characteristics. The bifurcation behavior is analyzed by using central manifold theorem through the way of phase trace, Poincare map, time history, power spectrum diagram and laypunov exponent & response under road stochastic excitation is obtained by computer simulation. After simulation it is revealed that the

chaos is   occurred in nonlinear suspension system.[7]

IX. CONCLUSION

By the literature review it is seen that compare with passive suspension system active suspension system can give the better handling & road holding characteristics. In earlier recherchés linear parameters of suspensions were considered but in practice the suspension parameters behaves nonlinear characteristic. So it is important to consider the nonlinearities of suspension system while designing the suspension system. This behavior of suspensions system is studied with half car model for better understanding of nonlinearities of suspension parameter because it can elaborate more detail than that of the study quarter car model.

REFERENCES:

[1] C Kim & P I Ro , "*Sliding model controller for vehicle active suspension system with nonlinearities*" IMechE1998

[2] Javad Marzbarad, Goodarz Ahmadi, Hassan Zohoor, Yousef Hojjat *Stochastic optimal preview control of vehicle suspension,* Journal of Sound and Vibration 275 (2004) 973–990

[3] Q. Zhu and M Ishitobi *Chaos and bifurcations in a nonlinear vehicle model*, Journal of Sound and Vibration 275 (2004) 1136–1146

[4] L.V.V. Gopala Rao, S. Narayanan *Preview control of random response of a half-car vehicle model traversing rough road,* Journal of Sound and Vibration 310 (2008) 352–365

[5] H Du and N Zahang presented work on *Constrained H∞ control of active suspension for a half-car model with a time delay in control.,* Proceedings of the Institution of Mechanical Engineering
Engineers, Part D: Journal of Automobile

[6] Mohammad Mehadi Fateh and Seyed Sina Alavi worked on the *Impedance control of an active suspension system.,* Mechatronics 19 (2009) 134–140, Elsevier.

[7] Yung sheng, Guang-qiang, Xian-jie Meng *Study on bifurcation, chaos and chaotic control of vehicle suspension with nonlinearities under road excitation.*

An Integrated and flexible software system for effective Teaching - learning process in Engineering Education

Brahmananda S.H.[1], Dr. V. Cyril Raj[2]

*(Department of Computer Science, Research Scholar Dr MGR Research & Educational Institute, Chennai
** (Head of the Department of Computer Science, Dr MGR Research & Educational Institute, Chennai

ABSTRACT

The aim of this research was to build and investigate the effect of an Effective learning tool in the present day teaching and learning process in Engineering education based on problem centered approach with four phases of effective instruction stated by Dr. David Merrill. The teaching-learning process is then monitored by the software system, which provides session by session feedback to the teachers, students and the supervisors, and the knowledge based system which facilitates the over-all process efficiently to all the stake-holders of the institution. The course feedback, Teachers feedback, Student feedback, Supervisors feedback and Result Analysis are delivered for corrective measures. Survey findings and post-intervention assessment outcomes were used to assess the student's and teacher's perception of their goals, satisfaction, motivation, and performance.

Our findings suggest that the teachers and students perceived high level of motivation in attaining their goals, and the students who undertook learning with this methodology had higher mean performance test scores.

Keywords - *Effective learning tool, First principles of Instruction, Motivation*

I. INTRODUCTION

Many Teachers in Engineering colleges are not properly trained to become effective instructors. Many of them never receive formal training in how students learn, what difficulties they encounter in learning, how to address student learning problems or even how to present material effectively to students. Teaching is broadly conceived as the dissemination of content knowledge to students. They use their personal learning experiences as the basis for selecting teaching methods, addressing only one learning style, the preferred learning style of the teacher.

Teacher should seek to affect learning not simply by presenting information for student absorption, but rather by working as guide, motivator and participant with the students. It is important that students are aware of this. It works for the professors as a real class demonstration of the idea that "teaching is figuring out what students know and then helping them make connections between new information and prior knowledge" (Cross and Steadman, 1996). Class should develop a student's critical reading and thinking skills. While students can achieve content mastery through lectures and reading assignments, knowledge that is constructed by teacher and student through cooperative efforts, such as discussion groups and debate, is more likely to promote analytical skills.

According to Bloom's Taxonomy teacher and student must have the shared goal of exploring material to enhance critical thinking (analytical skills). Improvement in organized writing indicates how well these skills are acquired. Students should have the ability to take in data (read), compare and contrast information in order to break into components (analyze), reorganize the components (synthesize), in order to express a new or individualized idea (formulate a thesis). They should then be able to compare and contrast various theses in order to find the best/most useful one (evaluation and application). When students master these skills they become effective learners in any field.

To match teaching style with learning style it is essential to know how college students learn. Though Bloom taxonomy is a useful tool, it is an outline not a detailed plan for the college classroom. Anthony Grasha's integrated model of teaching and learning (Grasha 1996), which was developed under the influence of William Perry (Perry, 1970). Grasha and Perry are more immediately useful tools than Bloom, because they are more adaptable and focus upon the learning of college level students in clear and practical terms.

Perry views the central experience of a college education as the student encounter with the multiplicity of ideas and opinions that constitute the body of knowledge. Perry empirically documents the process and demonstrates how the instructor can expect to encounter actual student learning. Understanding the cognitive skills of students is Perry's first principle in elevating them to a higher functioning level.

Perry's second stage of student learning is multiplism. Students encounter a great deal of uncertainty at this phase. The normally attentive college student encounters multiple answers for every question, which tests previous notions about the certainty of knowledge and threatens long-standing beliefs. As a result, puzzled by the apparent lack of standards, students either see all ideas as equally valid or equally biased, becoming suspicious of the truth of any evidence or authority. Perry found that this could cause students to avoid a thorough consideration of alternative views and to develop opinions largely on the basis of whim or personal belief (Culver and Hackos, 1982).

Perry noted that students do not advance through the dualistic to the multiplistic to the relativistic stages and achieve a real synthesis of knowledge until they can make a commitment to an idea or value that affirms their own identity. Commitment entails the realization that all ideas and dreams are fallible, changeable and eventually in need of reevaluation. In the end, a true commitment to knowledge results in the realization that all opinions and value may change. Furthermore, Perry clearly articulates, unlike Bloom, that knowing is an intimate engagement not a detached encounter. The well-prepared teacher must realize the intimacy of the teaching/learning experience and the fostering commitment in students entails changing student behaviour.

If students only master content, they have attained only the most rudimentary stage of learning, so teachers should learn to motivate students to improve their content acquisition, transforming the students from passive receptors of knowledge to active participants in the learning process. The key word is **active**. Learning that is active focuses on involving the students more directly in the learning process. It moves away from an emphasis upon the content to a focus upon developing student's skills to encounter the material. It shifts the responsibility for learning to the students and away from the teacher. The process can only be successful by modifying the preconception that the benefits of a college course accrue only within the walls of the classroom. Students must be made responsible for their learning at all times by making students accountable. Accountability is the vehicle that moves students to work to change their behavior outside of class, saving time for in class activities, which lead them to become more effective learners.

Michaelsen's team-based methods modify student behaviour by employing a technique called the Readiness Assurance Process. The Readiness Assurance Process initiates student accountability by informing students in the very first moments of a class about the objectives and the organizational framework that is being used to achieve class goals. This information empowers students to adapt their personal learning strategies to the class plan, reinforcing the idea of personal responsibility for the work at hand. In a typical college course, extra-class readings are a part of the class plan. With the Readiness Assurance Process students are tested on the concepts introduced by the readings at the start of each new class segment or lesson. Individual students initially take a test (Michaelsen recommends multiple choice tests) on the assigned readings followed immediately by the team attempting the same test as a group. The theory is to add to the accountability students normally have to the instructor in their personal work by making each student responsible to the other members of the team as well. Students are also given formal opportunities to evaluate team members. The principle is that peers are more aware of the efforts of their fellow students and that social pressure is a significant and more pervasive motivating force for students than the threat of the professor's grade alone.

The lesson learned from the Readiness Assurance Process is that strict accountability standards and peer review are powerful methods to modify stubborn student's behaviours. Frequent and timely feedback reinforces student responsibility and promotes effective learning.

All assignments, such as essays or exams or the Readiness Assessment Tests, must be structured in a fashion compatible with student intellectual levels and student learning styles (Michaelsen, Knight and Fink, 2002). They must include clear instructions on how students are to perform. Recall the discussion regarding student assessment above. Again, this applies to whatever teaching style the instructor uses. A second type of reinforcement is the creation in students of the expectation that their accountability is constant, that their learning will progress when they are prepared to progress, and that they will be held accountable in every class. Team based learning works well in this regard because it requires the students to produce a measurable product for every activity, and the team format can be monitored at every stage.

Maintaining accountability in students promotes responsibility among team members, a useful social skill, which enables students to work effectively with others. By working with others on a regular basis, students encounter different ideas and approaches, enhancing their ability to distinguish among multiple ideas. This is Perry's fourth level of knowing, critical thought, which every college instructor desires. It is "…deliberate, conscious thought or reflection that is desired toward accomplishing some goal …It has some purpose such as solving problems, making decisions, or applying information to our lives…(Grasha, 1996)." It is reasoned thought in that it enables one to consider a broad range of information relevant to an issue and then to develop an informed conclusion. And "critical thinking evaluates in a constructive manner more than one side of an issue as well as the positive and negative attributes of a situation (Michaelsen, Knight and Fink, 2002)."

According to Dr. M. David Merrill's First Principles of Instruction. a) Learning is promoted when learners are engaged in solving real world problems. b) Learning is promoted when existing knowledge is activated as a foundation for new knowledge. c) Learning is promoted when new knowledge is demonstrated to the learner. d) Learning is promoted when new knowledge is applied by the learner. e) Learning is promoted when new knowledge is integrated into the learner's world.

II. IMPLEMENTATION

The effective learning tool with the above described characteristics is being implemented using the Knowledge based systems; an integrated and flexible concept processing system is developed.

The contents of the course and the number of sessions will be prescribed by the University, and accordingly the lesson-plan will be designed by the instructor by taking into considerations of all the recommendations done based on the work of the Educational psychologists and the academicians which is then approved by the authorities of the institution, the knowledge based system helps the instructor in designing

the lesson plan, pedagogy materials, course objectives, monitors the process of teaching- learning.

The effort of building such a system is being done taking due care that the teachers and the students will have a smooth sail throughout the process without any extra work in achieving their goals in using this system.

Studies have shown benefits in the temporal association of visual and verbal information, where presenting visual and verbal sources at the same time leads to better learning than presenting them at different times (Mayer & Anderson, 1992; Mayer, Moreno, Boire, & Vagge, 1999). Benefits have also been found for spatial association, where learning is supported by placing visual and verbal materials in close physical proximity or integrating them into a single, combined representation (Hegarty & Just, 1993 Moreno & Mayer, 1999). One proposed rationale for these benefits is that temporal/spatial coordination reduces cognitive load demands associated with working memory maintenance and visual search (Mayer, 2001). The reduction in cognitive effort needed to find and maintain multiple sources of information allows students to engage in deeper processing. Visual-Verbal knowledge Integration (Multimedia) class sessions compared to the 19th century methodology of black-board classroom sessions have critical and powerful effects on learning.

The software aids in building an effective learning tool for both the teachers and students in accomplishing their goals, it facilitates the on-going sessions to all the students enrolled, it monitors the students session, It interacts with the students during session exercises individually/group, If a particular student fails in answering the question, instead of giving an immediate feedback on errors uses the model of desired performance. The model of desired performance refers to the behaviours or performance we desire students to achieve. In cognitive Tutors the model of desired performance is implemented as a set of production rules representing target skills in a specific domain. The model of desired performance plays a diagnostic role in intelligent tutor systems. When student behaviour is consistent with the model of desired performance, the system does not intervene. However, if student behaviour is inconsistent with the model of desired performance, the system intervenes with feedback so as to guide students toward performance that is consistent with the model.

Currently, feedback in Cognitive Tutors is based on what is broadly referred to as an expert model. An expert model feedback is structured so as to lead students toward expert-like performance. The tutor intervenes as soon as students deviate from a solution path. An alternative model that could serve as the basis for feedback is the assumption that someone with general skills facing a novel problem is still likely to make errors. Recognizing this possibility, the software incorporates error detection and error correction activities as part of the task. Feedback based on such a model would support the student in both the generative and evaluative aspects of a skill, while preventing unproductive floundering, goes a step further by providing the necessary contents, definitions, explanations so that the student understands all the related concepts and then helps the students in correcting their mistakes. It evaluates the students, monitors, motivates, encourages, reminds, if necessary for the stubborn students warns about the negative consequences of not attaining the long term goals which was agreed upon before/during the registration of the course. It provides feedbacks, maintains the results database of each session, it does the result analysis session-by-session helping in continuous evaluation.

A knowledge-based system is a computer system that is programmed to imitate human problem-solving by means of artificial intelligence and reference to a database of knowledge on a particular subject.

To be more specific, knowledge based system also called as expert systems are generally conceptualized as depicted in *Figure 1*. The user makes a consultation through the interface system (the communication hardware and also the software which defines the types of queries and formal language to be used) and the system questions the user through this same interface in order to obtain the essential information upon which a judgment is to be made. Behind this interface are two other sub-systems: - the knowledge base, made up of all the domain-specific knowledge that human experts use when solving that category of problems and - the inference engine, or system that performs the necessary reasoning and uses knowledge from the knowledge base in order to come to a decision with respect to the problem posed.
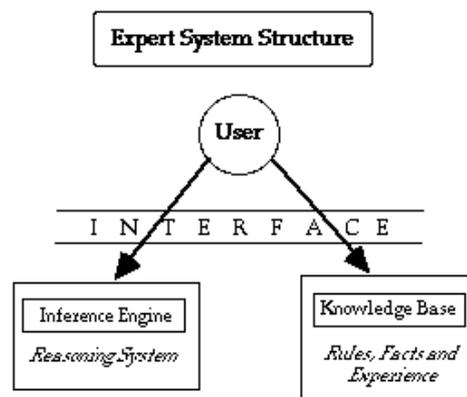
Figure 1 Knowledge based system

The knowledge based system here interacts with the management, Higher Officials, Supervisors, Teachers, Agents in fulfilling the vision and mission of the institution, and in attaining the desired goal of providing an efficient teaching-learning environment.- the knowledge base of the system.

The adaptive rule-based procedures or modules guides the teachers in preparing the course content, Identify key concepts, terms, and skills to be taught and learned based on the recommendations done by the Educational psychologists as mentioned earlier in making the teaching learning process more effective.
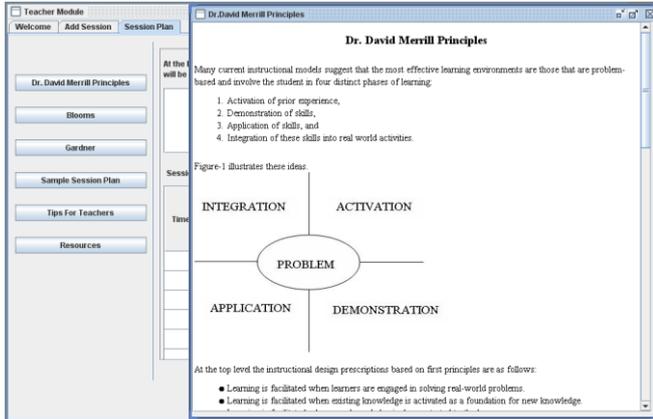
Fig 2: Assists the teacher while preparing the course content.

The teachers before and after preparing for each session will fulfill the criteria's by answering the queries such as: i) the course goal clearly mentioned to the teachers and the students. ii) Role of the teachers and students clearly mentioned. iii) Whether graphical representations (e.g. Graphs, figures) that illustrate key process and procedures used sufficiently where ever required. iv) Whenever possible, present the verbal description in an audio format rather than as written text. Whether the integration of audio and video pedagogy prepared which helps in understanding the subject more deeply. v) Whether pre-questions are prepared before the introduction of the new topic for knowing the readiness assessment test of the students. vi) Whether quizzes, multiple choice questions, exercises are prepared for each session and evaluated. vi) Encourage students to "think aloud" in speaking or writing their explanations as they study. vii) Encourage teachers to ask deep questions when teaching, and provide students with opportunities to answer deep questions to stimulate thought. viii) Feed backs at every stage are being monitored and corrective measures being taken at every stage by providing corrective feedback to the students and the teachers. – the inference engine of the system.

III. PROCEDURE AND DISCUSSION

As a case study, we chose C programming subject, which is a common subject for all Engineering students irrespective of which department the student belongs to (Students from Computer Science & Engineering, Information Science & Engineering, Electronics & Communication, Electrical & Electronics, Telecommunication Engineering, Mechanical Engineering, Civil Engineering were selected). Set of 57 students were taken in Experimental group and 53 students in Control group were chosen from the same Engineering stream (Information Science & Engineering), so as both the groups were divided only on the basis of their role numbers.

The Teachers in the Experimental group used the software which aided them in preparing their lesson plan, and to construct each and every session based on Dr. David Merrill's First principle of Instruction, After each session, the students attended the assignment/exercises which was posted by the teacher which was then monitored and supervised by the software on behalf of the teacher guiding the students in completing their assignment. The students

were reminded about their assignments by sending messages to their mobile phones which were registered at the beginning of the course by the software on behalf of the teacher, so that the students would feel that, teacher is monitoring and supervising their assignments online. The survey assessment showed that the 91% of students from the experimental group felt motivated and continued in doing their end-of-session exercises more regularly compared to only 42% of the students from the control group did their end-of-sessions regularly because students felt lack of supervision and motivation.
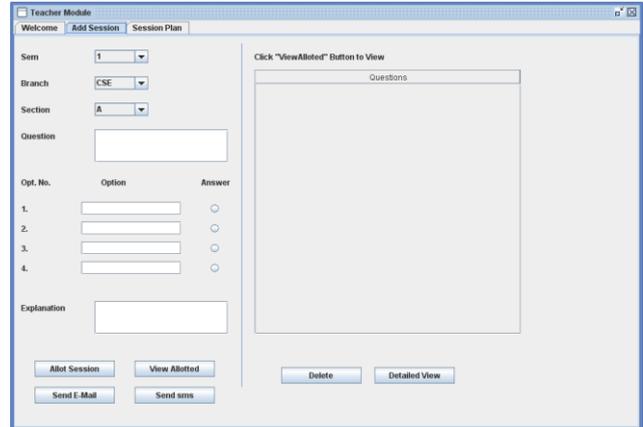
Fig 3. Assists both teachers and students in end-of-session exercises.

As every programs explained were demonstrated, compiled and executed in the experimental group, the students were able to understand much better unlike the black board teaching done in the control group. The students in the experimental group were able to develop and execute most of the programs assigned in the laboratories compared to the students in the control group, their motivation level while applying their knowledge to the problems assigned were also higher among the students in the experimental group.
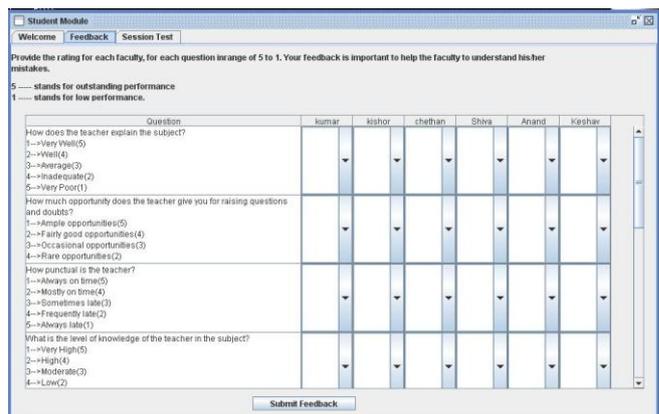
Fig 4. Assists in conducting and generating detailed report of the students feedback.

The result analysis is done with respect to the pre-mentioned goals at every stage and recommends the management for external motivation and corrective measures of each teacher and students of the institution. The results can be compared with other conventional

teaching methodologies being used and analyzed of the effectiveness of the learning tool.

Survey was conducted for both the teachers and the students with both the groups to explore the satisfaction, motivation, and learning orientation on a five point scale from strongly disagree to strongly agree, (from one to five), for item scoring. By the survey at the end of the session and throughout, the teachers following this integrated approach using this software scored very well compared to their counterparts who followed the conventional way of teaching.

Common test papers were given for both the groups set by a teacher (domain expert) not belonging to any of these experimental or Control group and the identity of the students were withheld to avoid bias while valuating the scripts. All the test scores were for Maximum of 25 marks.

Student's *t*-Test: Results

Group A (Experimental Group): Number of students= 57
16.0 16.0 17.0 17.0 17.0 17.0 17.0 18.0 18.0 18.0 19.0 19.0 19.0 19.0 20.0 20.0 20.0 21.0 21.0 21.0 21.0 21.0 21.0 21.0 22.0 22.0 22.0 22.0 22.0 23.0 23.0 23.0 23.0 23.0 23.0 23.0 23.0 23.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 24.0 25.0 25.0 25.0 25.0 25.0 25.0

Mean = 21.6

95% confidence interval for Mean: 20.66 thru 22.61

Standard Deviation = 2.70

Hi = 25.0 Low = 16.0

Median = 22.0

Average Absolute Deviation from Median = 2.23

Group B (Control Group): Number of students= 53
7.00 7.00 9.00 9.00 10.0 11.0 11.0 11.0 11.0 12.0 12.0 12.0 12.0 12.0 13.0 13.0 14.0 14.0 14.0 15.0 15.0 16.0 16.0 16.0 16.0 16.0 16.0 17.0 17.0 17.0 17.0 17.0 18.0 18.0 18.0 19.0 20.0 20.0 20.0 20.0 20.0 21.0 21.0 21.0 21.0 21.0 21.0 23.0 23.0 23.0 23.0 24.0 24.0

Mean = 16.3

95% confidence interval for Mean: 15.29 thru 17.31

Standard Deviation = 4.56

Hi = 24.0 Low = 7.00

Median = 16.0

Average Absolute Deviation from Median = 3.77

The results of an unpaired t-test performed showed

t= 7.52

sdev= 3.71

degrees of freedom =108

The probability of this result, assuming the null hypothesis, is less than .0001The probability of this result, assuming the null hypothesis, is less than .0001 by conventional criteria, this difference is considered to be extremely statistically significant.

IV. CONCLUSION

Our work showed that both Teachers and Students were highly motivated and encouraged while using this integrated and flexible software system which consistently held them in following the agreed upon goals throughout the course unlike where after going through a workshop or seminars the teachers motivation gets drained off slowly. This is of significance to engineering educators which helps in taking the students from passive receptors of knowledge to active participants in the learning process without much burden on the teachers, though this approach showed significant results in both motivation and remarkable performance of the students in programming subjects. This approach is to be extended for other engineering subjects also and investigated for more conclusive results.

REFERENCES

[1] Investigating the Effect of 3D Simulation-Based learning on the Motivation and Performance of Engineering Students. Caroline Koh, Hock Soon Tan, Kim Cheng Tan, Linda Fang, Fook Meng Fong, Dominic Kan, Sau Lin Lye, and May Lin Wee. Journal of Engineering Education July 2010.

[2] First Principles of Instruction Dr. M. David Merrill. ETR&D, Vol 50, No. 3,2002, pp43-59 ISSN 1042-1629

[3] Intelligent Learning Environments. P.Dillenbourg, M.Hilario, P.Mendelsohn, D.Schneider, B.Borcic.

[4] Creating the Teaching Professor: Guiding Graduate Students to Become Effective Teachers. Ronald J.Weber, Ann Gabbert, Joanne Kropp, and Patrick Pynes. The Journal of Scholarship of Teaching and Learning, Vol. 7, No.1, May 2007, pp.45-63.

[5] Bloom, Benjamin S. (1980). All Our Children Learning. New York: McGraw-Hill.

[6] John R.Anderson, Kevin A Gluck, Carnegie Mellon University, Psychology Department, Piisburgh, PA 15213, What role do cognitive architectures play in intelligent tutoring systems?

[7] Fostering the Intelligent Novice: Learning From Errors with Metacognitive Tutoring. Santhosh A. Mathan, Kenneth R. Koedinger, Educational Psychologist, 40(4), 257-265

[8] Grasha, A.F. (1996). Teaching with Style. Pittsburgh: Alliance Publishers.

[9] Tutorial on agent-based modeling and simulation Part-2: How to model with Agents. Charles M.Macal, Michael J. North Proceedings of the 2006 Winter Simulation Conference

[10] Weber, G (1999). Adaptive learning systems in the World Wide Web. In User Modeling; Proceedings of the Seventh International Conference UM'99(pp. 371-378). Springer.

[11] The Shared Expertise Model for Teaching Interactive Design Assistants. Tomasz Dybala, Gheorhe Tecuci, Hadi Rezazad. Engg Applic. Atrif. Intell vol.9 No.6 pp. 611-626. 1996

[12] Cypher.A (Ed), Watch What I Do: Programming by Demonstration. MIT Press, Cambridge, MA, 1993.

Prediction of Phishing Websites Using Optimization Techniques

R.Sumathi[1] and Mr.R.Vidhya Prakash[2]

[1]M.E Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India

[2]Assistant Professor, Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India

ABSTRACT

Phishing website is a fraudulent attempt usually made through email, to steal personal information. Phishing emails usually appear to come from a well-known organization and ask for personal information such as credit card number, social security number, account number or password. Often times phishing attempts appear to come from sites, services and companies with which do not even have an account. This paper presents a novel approach to overcome the difficulty and complexity in predicting and detecting phishing website. In existing system they proposed an intelligent resilient and effective model that is based on using association and classification Data Mining algorithms. They implemented PART classification algorithm and techniques to extract the phishing data training sets criteria to classify their legitimacy. In the proposed system, we implement the PSO algorithm for predicting Phishing Websites. In this project, we present novel approach to overcome the 'fuzziness' in the phishing website assessment and propose an intelligent resilient and effective model for phishing websites. The experimental results demonstrated the feasibility of using Association and Classification techniques and PSO real applications and its better performance.

KEYWORDS:- APRIORI, ASSOCIATION, CLASSIFICATION, DATA MINING, FUZZY LOGIC, PHISHING, PSO, RISK ASSESSMENT.

1. INTRODUCTION

Phishing is an e-mail fraud method in which the perpetrator sends out legitimate-looking email in an attempt to gather personal and financial information from recipients. Phishing is similar to fishing in a lake, but instead of trying to capture fish, phishers attempt to steal personal information. They send out e-mails that appear to come from legitimate websites such as eBay, PayPal, or other banking institutions. The e-mails state that information needs to be updated or validated and ask that enter username and password, after clicking a link included in the e-mail. Some e-mails will ask that to enter even more information, such as full name, address, phone number, social security number, and credit card number. However, even if we visit the false website and just enter username and password, the phisher may be able to gain access to more information by just logging in to account. The word phishing from the phrase "website phishing" is a variation on the word "fishing". The idea is that bait is thrown out with the hopes that a user will grab it and bite into it just like the fish. The motivation behind this study is to create a resilient and effective method that uses Data Mining algorithms and tools to detect phishing websites in an Artificial Intelligent technique. An Optimization Technique can be very useful in predicting phishing websites. It can give us answers about what are the most important phishing website characteristics and indicators and how they relate with each other. Comparing

between different Data Mining Optimization methods and techniques is also a goal of this investigation. The paper is organized as follows: Section A presents the literature review, Section B shows data mining phishing approach, Section C shows the theory and methodology of the research, Section D shows the utilization of the DM classification techniques, Section III reveals the conclusions and future work.

2. RELATED WORKS
2.1. Literature Review
A report by Gartner estimated the costs at $1,244 per victim, an increase over the $257 they cited in a 2004 report [1]. In 2007, Moore and Clayton estimated the number of phishing victims by examining web server logs. They estimated that 311,449 people fall for phishing scams annually, costing around 350 million dollars [2]. There are several promising defending approaches to this problem reported earlier.

One approach is to stop phishing at the email level [3], since most current phishing attacks use broadcast email (spam) to lure victims to a phishing website. Another approach is to use security toolbars. The phishing filter in IE7 [4] is a toolbar approach with more features such as blocking the user's activity with a detected phishing site. A third approach is to visually differentiate the phishing sites from the spoofed legitimate sites. Dynamic Security Skins [5] proposes to use a randomly generated visual hash to customize the browser window or web form elements to indicate the successfully authenticated sites. A fourth approach is two-factor authentication , which ensures that the user not only knows a secret but also presents a security token [6]. Many industrial anti phishing products use toolbars in Web browsers, but some Researchers have shown that security tool bars don't effectively prevent phishing attacks. Another approach is to employ certification, e.g., Microsoft spam privacy [7]. A variant of web credential is to use a database or list published by a trusted party, where known phishing web sites are blacklisted. The weaknesses of this approach are its poor scalability and its timeliness. The newest version of Microsoft's Internet Explorer supports Extended Validation (EV) certificates, coloring the URL bar green and displaying the name of the company. However, a recent study found that EV certificates did not make users less fall for phishing attacks [8].

2.2. Phishing Data Mining Approach

2.2.1. Phishing Characteristics and Indicators
There are many characteristics and indicators that can distinguish the original legitimate website from the phishing one. We managed to gather 27 phishing features and indicators and clustered them into six Criteria (URL & Domain Identity, Security & Encryption, Source Code & Java script, Page Style & Contents, Web Address Bar and Social Human Factor), and each criteria has its own phishing components. The full list is shown in table I which is used later on our analysis and methodology study.

2.2.2. Why use Data Mining?
DM is the process of searching through large amounts of data and picking out relevant information. It has been described as "the nontrivial extraction of implicit, previously unknown, and potentially useful information from large data sets [9]. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions [10].

3. THEORY AND METHODOLOGY
3.1. Data Mining Techniques
The approach described here is to apply data mining algorithms to assess e-banking phishing website risk on the 27 characteristics and factors which stamp the forged website. We utilized data mining classification and association rule approaches in our new phishing website detection model as shown in figure I to find significant patterns of phishing characteristic or factors in the e-banking phishing website archive data. Particularly, we used a number of different existing data mining association and classification techniques.Including JRip, PART [II], PRISM [12] and C4.5 [13], CBA [14], MCAR [15] algorithms to learn and to compare the relationships of the different phishing classification features and rules. The experiments of C4.5, RIPPER, PART and PRISM algorithms were conducted using the *WEKA* software 16]. CBA and MCAR experiments were conducted using an implementation provided by the authors of [14], [15]. We used two web access archives, one from APWG archive [17] and one from Phishtank archive [16]. We managed to extract the whole 27 phishing security features and clustered them to its 6 corresponding criteria as mentioned before in table 1.

3.2. Website Phishing Training Data Sets
Two publicly available datasets were used to test our implementation: the "phishtank" from the phishtank.com [16] which is considered one of the primary phishing-report collates both the 2007 and 2008 collections. The PhishTank database records the URL for the suspected website that has been reported, the time of that report, and sometimes further detail such as the screenshots of the website, and is publicly available. The Anti Phishing Working Group (APWG) which maintains a "Phishing Archive" describing phishing attacks dating back to September 2007 [3]. A data set of 1006 phishing, suspicious and legitimate websites is used in the study (412 row phishing websites, 288 rows suspicious and 306 row of real websites for the legitimate portion of the data set). In addition, 27 features are used to train and test the classifiers. We used a series of short scripts to programmatically extract the above features, and store these in an excel sheet for quick reference.

Table.1. Main Phishing Indicators with its Criteria

Criteria	N	Phishing Indicators
URL & Domain Identity	1	Using IP address
	2	Abnormal Request URL
	3	Abnormal URL of anchor
	4	Abnormal DNS record
	5	Abnormal URL
Security & Encryption	1	Using SSL certificate(Padlock Icon)
	2	Certificate Authority
	3	Abnormal Cookie
	4	Distinguished name certificate
Source code & Java Script	1	Redirect pages
	2	Straddling Attack
	3	Phanning Attack
	4	OnMouseOver to hide the link
	5	Server Form Handler (SFH)
Page style & Contents	1	Spelling errors
	2	Copying Websites
	3	Using forms with *Submit* button
	4	Using pop-ups Window
	5	Disabling right click
Web Address Bar	1	Long URL address
	2	Replacing similar char for URL
	3	Adding a Prefix or Suffix
	4	Using the @ symbol to confuse
	5	Using hexadecimal char codes
Social Human Factor	1	Emphasis on Security
	2	Public generic salutation
	3	Buying time to access accounts

4. DM Classification Techniques

4.1. Associative Classification Algorithms

The practical part of this comparative study utilizes six different common OM classification algorithms (C4.5, JRip, PART, PRISM, CBA and MCAR). Our choice of these methods is based on the different strategies they used in learning rules from data sets. The C4.5 algorithm [13] employs divide and conquer approach, and the RIPPER algorithm uses separate and conquer approach. The choice of PART algorithm is based on the fact that it combines both approaches to generate a set of rules. PRISM is a classification rule which can only deal with nominal attributes and doesn't do any pruning. CBA algorithm employs association rule mining [14] to learn the classifier and then adds a pruning and prediction steps. Finally, MCAR algorithm consists of two phases: rules generation and a classifier builder. In the first phase, MCAR scans the training data set to discover frequent single items, and then recursively combines the items generated to produce items involving more attributes. MCAR then generates ranks and stores the rules. In the second phase, the rules are used to generate a classifier by considering their effectiveness on the training data set [IS].
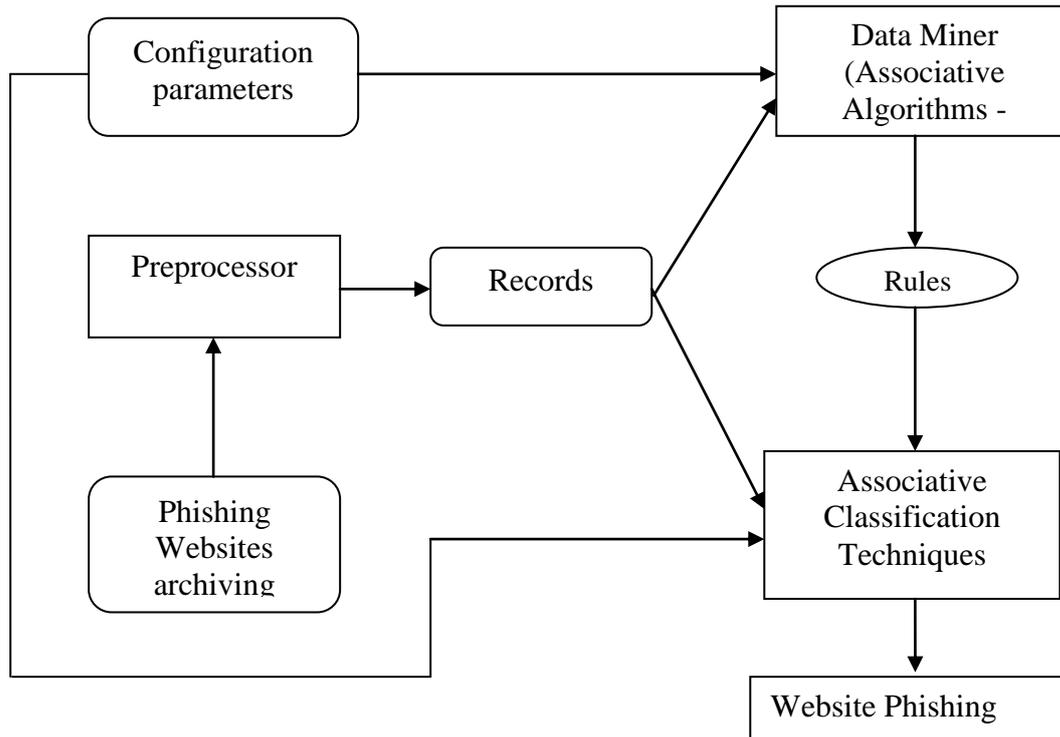
Figure.1 AC Model for Detecting Phishing

4.2. MCAR Phishing Model Approach Associative

Classification is a special case of association rule mining in which only the class attribute is considered in the rule's right-hand-side 11(consequent), for example A, B -) Y, Then A, B must be input items attributes and Y must be the output class attribute. The attribute values for all our input items which represent the six ebanking phishing features and criteria ranged between three fuzzy set values (Genuine, Doubtful and Legitimate) which we measured before in our previous paper using Fuzzy Logic [17] taking into consideration all the input fuzzy variables for all criteria different components as shown in Table I. The output class attribute of our ebanking phishing website rate is one of these values *(Very Legitimate, Legitimate, Suspicious, Phishy or Very Phishy).* Example of the training phishing data sets to be classified is shown in Table 2. To derive a set of class association rules from the training data set, it must satisfy certain user-constraints,i.e support and confidence thresholds. Generally, in association rule mining, any item that passes *MinSupp* is known as a frequent item. We recorded the prediction accuracy and the number of rules generated by the classification algorithms.

Table 2.Example of Training Phishing Data Sets

Row ID	URL	Security	Java	Style	Address	Social	Class/Phishing
1	G	G	D	G	G	G	**Very Legitimate**
2	D	G	G	D	G	D	**Legitimate**
3	D	D	G	F	D	G	**Suspicious**

4	F	D	G	D	F	D	**Phishy**
5	D	F	F	D	F	F	**Very Phishy**
*	**G=Geniue**			**D=Doubtful**			**F=Fraud**

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

Table 3 Results From Weka four Classifiers

	C4.5	P.A.R.T	JRip	PRISM
Test Mode	10 FOLD CROSS VALIDATION			
Attributes	URL Domain Identity Source Code & Java Web Address Bar		Security & Encryption Page Style & Contents Social Human Factor	
No. of Rules	57	38	14	155
Correct Classified	848 (84.2 %)	869 (86.3%)	818 (81.3 %)	855 (84.9 %)
InCorrect Classified	158 (15.7%)	137 (13.6%)	188 (18.8%)	141 (14.0%)
No. of instances	1006	1006	1006	1006

Table 4.Results from CBA and MCAR Classifiers

	CBA	MCAR
Num of Test Case	1006	1006
Correct Prediction	873	886
Error rate	13.452%	12.622%
Min Sup	20.000%	20.000%
Min Conf	100.000%	100.000%
Number of rules	15	22

4.3. Particle Swarm Optimization

Particle Swarm Optimization is a population based heuristic optimization algorithm inspired by social behavior of birds flocking or fish schooling. Each particle is treated as a point in a D-dimensional space. Initially, N particles are uniformly distributed in the solution space. The particles in PSO fly through the search space with a certain velocity, and change their position dynamically in the hope of reaching the food source, the destination. Therefore, position and velocity are two important parameters in the PSO algorithm.

Each particle keeps track of the best position it has encountered during its travel, and the best position traveled by the swarm of particles. The best position traveled by a particle is called the *local best position*, and the best position traveled by the swarm is called the *global best position*. At the end of each iteration, the particles calculate their next velocity, and update their positions based on the calculated velocity.

Basic algorithm for PSO:

1. Initialize

    (a) Set constants k_{max}, c_1, c_2.

    (b) Randomly initialize particle positions $x^i_o \in D$ in IR^n for $i = 1, \dots, p$.

    (c) Randomly initialize particle velocities $0 \le v^i_o \le v^{max}_o$ for $i = 1, \dots, p$.

    (d) Set $k = 1$

2. Optimize

    (a) Evaluate function value f^i_k using design space coordinates x^i_k.

    (b) If $f^i_k \le f^i_{best}$ then $f^i_{best} = f^i_k$, $p^i_k = x^i_k$.

    (c) If $f^i_k \le f^g_{best}$ then $f^g_{best} = f^i_k$, $p^g_k = x^i_k$.

    (d) If stopping condition is satisfied then goto 3.

    (e) Update all particle velocities v^i_k for $i = 1, \dots, p$

    (f) Update all particle positions x^i_k for $i = 1, \dots, p$

    (g) Increment k.

    (h) Go to 2(a)

3. Terminate

Figure 2. Flow Diagram of PSO-SVR forecasting model

Initialize the parameters of PSO

Randomly initialize x_i and v_i

For each particle, do SVR training
Then calculate particle fitness $f(x_i)$

Update the positions and velocities

Stop criteria

No

Yes

Choose the best parameters
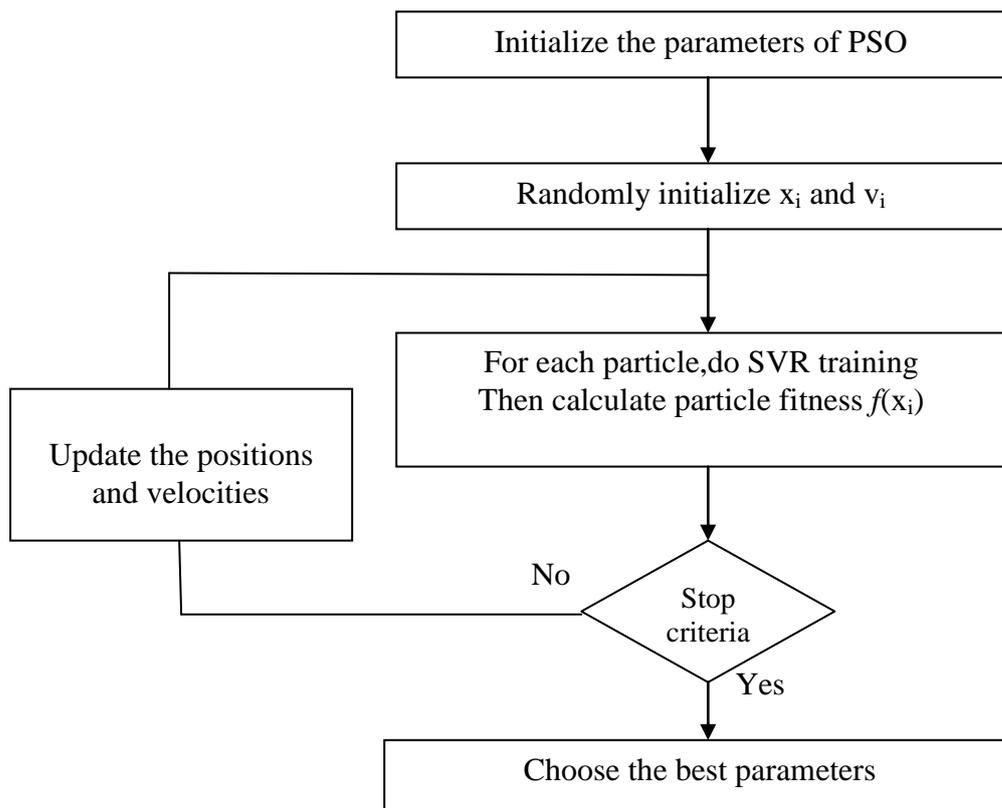
Table 5. Results From MCAR and PSO four Classifiers

	MCAR	PSO
Num of Test Case	1006	1006
Correct Prediction	886	934
Error rate	12.622%	9.544%
Min Sup	20.000%	20.000%
Min Conf	100.000%	100.000%
Number of rules	22	27

5. CONCLUSIONS

The Partial swarm optimization data mining phishing website model showed the significance importance of the phishing website two criteria's (URL & Domain Identity) and (Security & Encryption) with insignificant trivial influence of some other criteria like 'Page Style & content' and 'Social Human Factor' in the final phishing rate, which can help us in building website phishing detection system. The experiments indicate that Partical swarm optimization technique is highly competitive when compared with other traditional classifications in term of prediction accuracy and efficiency.

REFERENCES

[I]   GARTNE R, INC. Gartner Says Number of Phishing Emails Sent to U.S. Adults Nearly Doubles in Just Two Years, http ://www.gartner.com/it/page.jsp?id=498245.November 9 2006 .

[2]   T. Moore and R. Clayton, "An empirical analysis of the current state of phishing attack and defense", In Proceedings of the Workshop on the Economics of Information Security (WEIS2007).

[3]   B. Adida, S. Hohenberger and R. Rivest , "Lightweight Encryption for Email," USENIX Steps to Reducing Unwanted Traffic on the Internet (SRUTI), 2005 .

[4]   T. Sharif, "Phishing Filter in IE7," http://blogs,msdn.com/ie/archive/2005 /09/09/463204,aspx,,2006,

[5]   R. Dhamija and J.D. Tygar, "The Battle against Phishing: Dynamic Security Skins," Proc , Syrnp. Usable Privacy and Security, 2005.

[6]   FDIC., "Putting an End to Account-Hijacking Identity Theft," fdic .gov/idtheftstudy/identity_theft.pdf, 2004 , [7] Microsoft, "microsoft,com/twc/privacv/spam", 2004

[8]   C. Jackson, D. Simon, D. Tan, and A. Barth, "An evaluation of extended validation and picture-in-picture phishing attacks". In Proceedings of the 2007 Usable Security. www. usablesecurity.orgipapers/jackson.pdf

[9]   Kantardzic and Mehmed. *"Data Mining: Concepts, Models, Methods, and Algorithms ,",* John Wiley & Sons. ISBN 0471228524. OCLC 50055336,2003.[10] U,M, Fayyad, "Mining Databases: Towards Algorithms for Discovery," Data Eng, Bull., vol. 21, no, 1, pp, 3948,1998.

[II]   I.H. Witten and E. Frank, "Data Mining : Practical machine learning tools and techniques", 2nd Edition,Morgan Kaufmann, San Francisco, CA, 2005 ,

[12]   J. Cendrowska., *"PRISM: An algorithm for inducingmodular rule",* International Journal of Man-Machine Studies (1987), Vo1.27, No.4, pp.349-370.

[13]   J. R. Quinlan, "Improved use of continuous attributes in c4.5", Journal of Artificial Intelligence Research, 4:7790,1996,

[14]   Bing Liu, Wynne Hsu, Yiming Ma, "Integrating Classification and Association Rule Mining."*Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98,Plenary Presentation),* New York, USA, 1998.

[15]   '1', Fadi, c.Peter and Y. Peng, *"MCAR: Multi-classClassification based on Association Rule",* IEEE International Conference on Computer Systems and Applications ,2005, pp. 127-133.

[16]   WEKA - University of Waikato, New Zealand, EN,2006: "Weka -Data Mining with Open Source Machine Learning Software in Java", 2006 ,

AUTHORS

Ms.R.Sumathi received B.E degree in Computer Science and Engineering from Avinashilingam University, Coimbatore and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. Her research interest includes Data mining and Computer Networks.

Mr.R.Vidhya Prakash received the M.E degree in Software Engineering from P.S.G College of Technology, Coimbatore and received the B.E degree in Computer Science and Engineering from M.P.Nachimuthu M.Jaganathan Engineering College under Anna University Chennai. He is currently working as Assistant Professor in Department of Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology,Coimbatore. He has presented papers in conferences. His main research interest is Software Engineering and Computer Networks.

GREEN BUILDINGS – ON THE MOVE

B. Sai Doondi[1], Sarath Chandra Kumar B[1], Dr. P Saha[2]

[1](students, Department of Civil engineering, K L University, Vaddeswaram, A.P.-522502, India)
[2](Assoc Professor, Department of Civil engineering, K L University, Vaddeswaram, A.P.-522502, India)

ASTRACT:
Green building education requires successful teamwork of students from different disciplines in order to solve challenging problems in construction and design of buildings. "sustainable building" is the design and construction of buildings using methods and materials that are resource efficient and that will not compromise the health of the environment or the associated health and well being of the building's occupants, construction workers, the general public, or future generations. Sustainable building involves the consideration of many issues, including land use, site impacts, indoor environment, energy and water use, solid waste, and lifecycle impacts of building materials. Making existing and new buildings is one of the most effective levers to meet the challenges of CO2 reduction in cities. This paper discusses the concept of green buildings which are buildings based on sustainable principles, designed, built, renovated, operated and reused in an ecological and resource efficient manner and also presents a case study of building which is designed and constructed based on concept of green building.

Keywords: Green buildings, building materials, recycle & reuse, conservation

1. INTRODUCTION

India is witnessing tremendous growth in infrastructure and construction development. The construction industry in India is one of the largest economic activities and is growing at an average rate of 9.5% as compared to the global average of 5%. As the sector is growing rapidly It has in turn lead to many hazardous problems such as depletion of non renewable resources, generation of consumer waste on large scale, deforestation etc. which created a need to introduce a new concept called Green building.

Green building means a high performance building property that considers and reduces its impact on the environment and human health. Green buildings are designed and constructed to maximize whole life-cycle performance, conserve resources, and enhance the comfort of their occupants. This is achieved by the smart use of technology such as fuel cells and solar heated water tanks, and by attention to natural design elements such as maximizing natural light and building orientation. The result is a highly efficient building that saves money, is aesthetically pleasing, and contributes to the comfort and productivity of its occupants.

The goal of making Green building can be achieved by
- Efficiently using energy, water and other non renewable resources,
- Reducing waste, pollution and other environmental impacts
- Use of environmental friendly materials etc.

2. GREEN BUILDING CONCEPTS AND DESIGN

Green building construction practices that helps in lessen environmental impacts and improve the energy performance of new Constructions by a few fundamental principles that constitute the IGBC(Indian green building Council) designation, namely:

- Site selection and architectural planning
- Water management
- Energy efficiency and renewable energy
- Waste management
- Indoor environmental quality

SITE SELECTION AND ARCHITECTURAL PLANNING

Accessibility to basic amenities:
The site should be selected such that it is near public transit and/ or household services and amenities that are accessible by safe, convenient pedestrian pathways. The basic amenities include school, bank/ATM, crèche, medical clinic/ hospital, pharmacy, grocery store, electrician/plumbing services, dhobi/laundry, fitness centre, post office, place of worship, restaurant, supermarket, playground, electricity/ water utility bills payment counter and other neighborhood- serving retail.

Soil Erosion Prevention & Control:
Evolve strategies to stockpile top soil and reuse later for landscaping purpose. Stockpiled soil can be donated to other sites for landscaping purpose.

Natural Topography and Landscape:
Avoid site disturbance by retaining the natural topography of the site and / or landscape at least 20% of the site area or meet the local regulation to extent possible. In sites which have fully grown trees, destruction is to be avoided. Also avoid developing paved surfaces on the site as much as possible.

Reduction of Heat Island Effect on Roof and Parking Area:
Reduce heat islands that are reducing thermal gradient differences between developed and undeveloped areas to minimize impact on the microclimate. Use high Albedo roofing material or heat resistant paint or china mosaic or white cement tiles or any other highly reflective materials over the roof to cover atleast 50% of the exposed roof area. Provide vegetation to cover atleast 50% of the exposed roof area and Plant shade-giving trees to cover atleast 75% of the open parking areas (or) install permanent roof to cover 75% of the parking areas.

Design for Differently Abled:
Ensure that the factory building is user-friendly for differently abled people. Appropriately designed preferred car parking spaces in areas which have easy access to the main entrance or closer to the lift. Uniformity in flooring level / ramps in the factory areas. Rest rooms (toilets) designed for differently abled people.

WATER MANAGEMENT
Rainwater Harvesting:
The design should incorporate rainwater harvesting to increase the ground water table or to reduce the usage of water through effective and appropriate rainwater management. Capture rainwater at least 50% from the roof and non roof for reuse. The design should also include flushing arrangement to let out impurities in the first few showers. Such pollutants and impurities include paper waste, leaves, bird droppings, dust, etc.

Low Flow Water Fixtures:
To minimize indoor water usage by installing efficient water fixtures. While selecting water fixtures, look for the flow-rates. The product catalogue or the brochure may detail the flow rates at various pressures..

Turf design:
Turf design is to limit such landscapes which consume large quantities of water. Select turf, plants, shrubs and trees which consume less water and are resilient to local climatic conditions.

Drought Tolerant Species:
Ensure that atleast 30% of the landscaped area is planted with drought tolerant species. Select species that are well-adapted to the site.

Management of Irrigation System:
Reduce the demand for irrigation water through water-efficient management techniques. Provide highly efficient irrigation systems incorporating features mentioned below:

- Provide a central shutoff valve for the irrigation system.
- Provide a moisture sensor controller.

- Install time based controller for the valves such that the evaporation loss is minimum and plant health is ensured.
- Any other innovative methods for watering.

The designer and the installer must work together and ensure the design performance of the system.

Grey water treatment
Calculate the wastewater volumes generated in the building. Design appropriately the capacity of the on-site wastewater treatment system. While designing the treatment system,ensure that the treated wastewater meets the required quality standards based on its purpose of application.Grey water is neither clean nor heavily soiled waste water that comes from clothes washers, bathroom, bathtub, wash basins, showers, kitchen sinks and dish washers. More specifically, it is the untreated waste water which has not come into contact with toilet waste. An onsite grey water treatment plant to treat at least 50% of grey water generated in the building should be installed.

Treated grey water for landscaping and flushing
The demand for fresh water should be reduced by using treated grey water as much as possible. At least 50% of water requirement for landscaping should be met by using treated grey water generated within the site. The treated grey water should also be used for meeting at least 50%flushing requirements. Separate water plumbing lines should be installed to carry treated grey water for flushing requirements.

Water metering
All the major water consuming areas should be fitted with systems to monitor their consumption so that probable water saving can be predicted. Some of the areas that require water meters are treated grey water consumption, landscape water consumption, rain water reuse, air conditioning cooling tower makeup, hot water consumption, swimming pools, water fountain, common car wash facilities, etc.

ENERGY EFFICIENCY AND RENEWABLE ENERGY
CFC-Free Equipment:
Avoid the use of CFC based refrigerants and ozone layer depleting gases which negatively
Impact the environment. Install HVAC equipment which does not use CFC based refrigerant.

Minimum Energy Performance:
Optimise energy efficiency for non-process use in the building to reduce environmental impacts from excessive energy consumption. Identify the materials and equipment available in the market and their properties with regard to energy performance. While selecting these material and equipment, consider their associated environmental impacts. Determine the applications where automatic controls can help in

energy savings. Obtain details of the controls and ensure proper installation.

HCFC Free / Low Impact HCFC Equipment:
Avoid the use of HCFC based refrigerants and ozone layer depleting gases which negatively impact the environment. Install fire suppression systems which do not contain CFCs, HCFCs, HFCs or Halons.

Energy Metering:
To encourage continuous monitoring and enhance the performance of buildings. Identify all the major energy and water consuming equipment and install systems to monitor their consumption Have separate meters for process and non process loads. Provide meters for of the following items like Energy meter for air-conditioning, internal lighting, external lighting, water consumption, water pumping for landscaping, etc.

On-Site Renewable Energy:
Promote self sufficiency in energy through renewable technological sources of renewable energy that can be considered under this credit include solar energy, wind energy, biomass, biogas etc. for on-site power generation and use within the building.

Eco-Friendly Captive Power Generation:
It is nothing but to reduce emission levels and their impacts on environment through the use of low emitting fuels or better equipment. Such as Use of bio fuels or non edible oils or any other non-fossil based fuel for captive power generation .Use diesel generator sets which are certified by Central Pollution Control Board (CPCB) for emissions and noise compliance. Use ISI rated generator sets/

Efficient luminaries & lighting power density:
After careful assessment of economic viability, energy efficient light fittings should be fitted. These include efficient tubular fluorescent light fittings with electronic ballasts, T5 lamps, compact fluorescent light fittings, light emitting diodes, etc. The installed light fittings should be at least three star rated under BEE labeling programme. The lighting power density should be maintained within limits.

WASTE MANAGEMENT:
Handling of Non-process Waste:
Ensure effective non-process waste management, post occupancy for recycling and safe disposal. Have a facility to segregate at least five of the following non-process waste generated in the building. Organic waste,Plastic,Paper,Paperboard,Glass,Metals,'e'waste,Lamps,Batteries,etc.

Waste reduction during construction:
The waste generated during construction should be segregated based on its utility and should be sent for recycling. This will reduce waste going to landfills. Avoid at least 50% of the waste generated during construction being sent to landfills and incinerators. Typical construction debris include broken bricks, steel bars, broken tiles, glass, wood paste, paint cans, cement bags, packing material, etc.

Materials with Recycled Content:
Encourage the use of products which contain recycled materials to reduce environmental impacts associated with the use of virgin materials. Some of the materials with recycled content are fly ash blocks, tiles, steel, glass, cement, false ceiling, aluminum and composite wood.

Local Materials:
Low VOC Ensure that at least 50% of the total building materials by cost used in the building are manufactured within a radius of 500 km. Use of building materials available locally should be maximized thereby minimizing the associated environmental impacts.

Materials:
Use of materials with low emissions should be encouraged so as to reduce adverse health impacts for building occupants. All the possible interior materials which can have high VOC content is to be listed and replaced with materials with no or low VOC content based on durability, performance and environmental characteristics.

INDOOR ENVIRONMENTAL QUALITY:
Tobacco Smoke Control:
Minimize exposure of non-smokers to the adverse health impacts arising due to passive Smoking. Prohibit smoking in common areas like corridors, lobby, lifts etc., Building should be designed to eliminate or minimize tobacco smoke pollution in the common areas.

Minimum Fresh Air Requirements:
Naturally conditioned buildings may consider having window openings to bring in the fresh air. In case of forced ventilation systems, fresh air can be pumped into the spaces. In areas where the fresh air temperatures are either too high or too low, consider treating such air using systems like geo-thermal, wind towers, earth tunnel cooling, direct / indirect evaporative cooling etc.,

Day-lighting
During design stage, the orientation of the building should be adjusted such that maximum day-lighting to all the spaces is achieved for most part of the day.

Exhaust systems:
Exhaust systems in bathrooms and kitchens should be adequately designed to maintain indoor air quality.

3. INDIAN GREEN BUILDING CODE (IGBC) GREEN HOMES RATING SYSTEM:
The council encourages builders, developers and owners to build green to enhance the economic and environmental performance of buildings. IGBC continuously works to provide tools that facilitate the adoption of green building practices in India. The concept of a rating would encourage designers to address these by design. IGBC has set up the Green

Building Core Committee to develop the rating programme. This committee comprised of key stakeholders including corporate, architects, consultants, developers, manufacturers and institutions. As a general guideline, individual owners can use the checklist 'Projects with Interiors' and developers & builders can use the checklist titled 'Projects without Interiors'. The threshold criteria for certification levels are as given in Table 1:

Table 1 IGBC Green Homes Certification Criteria

Certification level	Points for projects with interiors	Points for projects without interiors
certified	32-39	30-36
Silver	30-37	37-33
Gold	38-59	35-55
platinum	60-80	56-75

4. GREEN RENOVATION:

The green building guidelines, design process, and team approach developed for new construction can also apply to building renovations. Green renovation practices include use of natural design elements, such as increased day lighting; installation of Resource conserving materials and systems; recycling and reuse of construction and demolition waste; and maintenance of good indoor environmental quality during construction. While renovation often provides an opportunity to improve indoor-air-quality problems and upgrade HVAC, electrical, solid waste, and water equipment and systems, the process can also have extensive environmental impacts on building occupants. Primary impacts include indoor-air-quality contamination, waste, noise, and hazardous conditions. These can be addressed by scheduling work when occupants are away from a building, placing barriers between the renovation and occupied areas, and providing adequate ventilation. Of special concern during the renovation process is the increased risk of occupant exposure to hazardous waste. The most common types of hazardous waste include asbestos, lead paint, Plastic material ,mercury-containing fluorescent lamps. Proper handling, disposal, or recycling of these materials is critical to ensure compliance with federal environmental regulations. Hazardous materials that are moved off-site can bring long-term environmental liabilities to the building owner or operator when not disposed of properly.

5. CONCLUSION:

The principles and guidelines discussed imply that Green Buildings can have tremendous benefits, both tangible and intangible. The immediate and most tangible benefit is in the reduction in operating energy and water consumption right from day one, during the entire life cycle of the building. The energy savings could range from 25 – 30 % depending on the extent of green specifications. Other tangible savings would be reduction in first costs and enhanced asset value. Intangible benefits of Green Buildings include increasing productivity of occupants 'health, safety benefits and a green corporate image. The building studied showed that it scored a total of 57 points out of a possible 75 points and could possibly get Platinum rating (without interiors) under IGBC. If the future builders and developers continue to follow such green concepts and develop green housing sectors, the trend will lead to more energy efficient homes and protect the environment.

REFERENCES:

[1] IGBC Green Homes Rating System Ver 1.0, Abridged Reference Guide (2009), Confederation of Indian Industry, CII-Sohrabji Godrej Green Business Centre, Hyderabad.

[2] David Rodman and Nicholas Lenssen, "A Building Revolution: How Ecology and Health Concerns are Transforming Construction," Worldwatch *Paper 124* (March 1995), 41.

[3] IGBC HOMES, http://www.igbc.in/site/igbc/testigbc.jsp?desc=115890&event=115679

[4] Turner Wayne C., Doty Steve (2007). Energy Management Handbook, the Fairmont Press, Inc. ISBN 0-88173-532-6.

[5] Sustainable building technical manual, green building design, construction and operation. Produced by Public Technology Inc. US Green Building Council, part VI

Automatic Discovery and Recognition of Activities in Smart Environment

L.Karthika*, A.C.Sumathi**

*II M.E CSE, Sri Shakthi Institute Of Engineering and Technology, Anna University, Coimbatore.
** Assistant Professor, CSE, Sri Shakthi Institute Of Engineering and Technology, Anna University, Coimbatore.

ABSTRACT:

The machine learning and pervasive sensing technologies found in smart homes offers opportunities for providing health monitoring and assistance to individuals experiencing difficulties living independently at home. To monitor the functional health of smart home residents, there is a need to design technologies that recognize and track activities that people normally perform as part of their daily routines. The existing approaches are applied to activities that have been preselected and for which labeled training data are available. For this, an automated approach is proposed for activity tracking, which identifies frequent activities that naturally occur in an individual's routines. With this capability, the occurrences of regular activities are monitored and also can detect the changes in an individual's patterns.

Keywords—Activity recognition, Data mining, sequence mining, clustering, smart homes.

INTRODUCTION

A convergence of technologies in machine learning and pervasive computing as well as the increased accessibility o robust sensors and actuators has caused interest in the development of smart environments to emerge. Smart Environments can assist with valuable functions such as remote health monitoring and intervention. The need for the development of such technologies is underscored by the aging of the population, the cost of formal health care, and the importance that individuals place on remaining independent in their own homes. To function independently at home, individuals need to complete Activities of Daily Living (ADLs) [1] such as eating, dressing, cooking and drinking. Automating the recognition of activities is an important step toward monitoring the functional health of a smart home resident.

In response to this recognition need, several approaches to model and recognize are designed. The generally accepted approach is to model and recognize those activities that are

frequently used to measure the functional health of an individual [2]. However there is a number of difficulties in this approach. First, there is an assumption that each individual performs most, or all, standard ADLs in a consistent predefined manner in their home environments where they can be monitored. But this is certainly not always the case. In addition, the same individual might perform even the same activity indifferent ways, requiring methods that can also deal with intrasubject variability.

Second, tracking only preselected activities ignores the important insights that other activities can provide on the functional health of an individuals. This highlights the fact that it is important for a caregiver to recognize and monitor all activities that an individual regularly performs in their daily environments.

Third, to track a predefined list of activities, a significant amount of training data must be labeled and made available to the machine learning algorithm. Unfortunately, collecting and labeling such sensor data collected in a smart environment is an extremely time-consuming task.

In this paper, an unsupervised method is introduced for discovering and tracking activities in a smart environment that addresses these issues. This project based on the context of the CASAS Smart Home Project [9] by using the sensor data that are collected in the CASAS smart apartment testbed. The unsupervised nature of this model provides a more automated approach for activity recognition than offered by previous approaches.

Compared to traditional methods, for activity recognition which uses HMM model, our first approach is that discover the frequent activities by using the unique mining method, along with a clustering step to group discovered patterns. For the recognition step, an extended version of a hidden Markov model (HMM) is created to represent the activities and their variations, and to recognize those activities when they occur in the smart environment.

In the remainder of this paper, the approaches of this project are activity discovery, recognition, and tracking are

explained. In section II, these approaches are compared with the related work. In section III, discovery of activities using mining and clustering methods. Then this paper's main contribution in section IV describes how discovered activities can be recognized using the HMM model.

RELATED WORK

Activity recognition needs different approaches. These approaches differ accordingly to the type of sensor data, activity model and the methods to annotate the sample data.

A. Sensor Data

The different types of sensor information are effective for classifying different types of activities. Previously [3], [4] are collected the sensor information from the state-change sensors and RFID tags. Some researchers such as [5] processed the video to recognize the activities. While accessing the video is very complex.

B. Activity Models

The numbers of machine learning models are used for activity recognition. Naïve Bayes classifiers have been used with promising results for activity recognition [6]. In this approach, Hidden Morkov Model is employed to recognize the related activities from stream of sensor information.

C. Annotation Methods

An aspect of activity recognition, a method is used to annotate the sample data. Most of the researchers have been used the labeled information for activity recognition [4], [5]. This is not always the case.

Here a new mining method is introduced called, Discontinuous Varied-Order Sequential Mining (DVSM), which is able to find frequent pattern of activities that may be discontinuous and might have variability in the ordering. Intrasubject variability issue is addressed. Activity clustering is employed here to group the patterns. In the next step the Hidden Morkov Model is used to represent the activities and their variations and to recognize those activities. The architecture of the system is shown in Fig. 1.
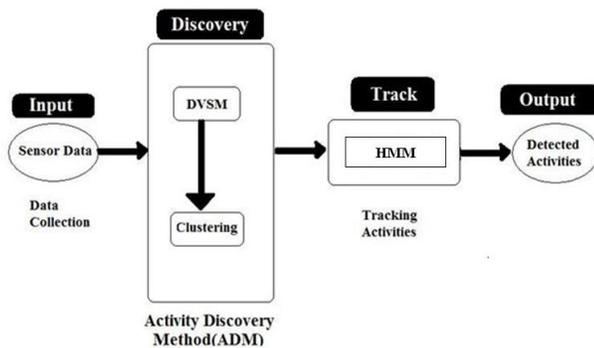
Figure.1   Main Components for discovering and tracking activities

DISCOVERING ACTIVITIES

The first step, the frequent and repeatable sequences of sensor events are considered, that comprise our smart environment's notion of an activity. By applying frequent sequential pattern mining techniques, with this contiguous events are identified, consistent sensor event sequences that might indicate an activity of interest. Ruotsalainen and Ala-Kleemola [7] introduce their Gais algorithm for detecting interleaved patterns using genetic algorithms, but this is a supervised learning approach that looks for matches to specific pattern templates. Given that sequential patterns are discoverd that may be discontinuous and have variability in the ordering, another possible approach is to cluster the sensor events.

The limitation of clustering algorithms for our problem is that all of the data points are don not clustered, but only those that are part of an activity sequence which is likely to occur frequently and with some degree of regularity or recognisability. So sequence mining and clustering algorithm is combined into an Activity Discovery Method (ADM) to identify frequent activities and cluster similar patterns together.

A. Discovering Frequent Discontinuous Sequences

DVSM is used to find discontinuous instances. This approach is different from frequent item set mining because the orders of items are considered as they occur in the data.

To find the Discontinuous Sequences from the input data D, DVSM first creates a reduced data set D_r containing the top most frequent events. Next, DVSM slides a window of size 2 across D_r to find patterns of length 2. DVSM extends the patterns discovered in the previous iteration by their prefix and suffix events, and will match the extended pattern against the already discovered patterns to see if it is a variation of a previous pattern, or if it is a new pattern.

Levenshtein (edit) distance [8] is used to define a similarity measure sim(A,B) between the two patterns. The edit distance, e(A,B), is the number of edits. The similarity measure is defined based on the edit distance as

$$sim(A,B) = 1 - \left(\frac{e(A,B)}{\max(|A|,|B|)} \right)$$

The continuity between component events, Γ_e, is defined or each two consecutive events in an instance. Then $\Gamma_e(e')$, the event continuity for e' is defined as

$$\Gamma_e(e') = \frac{1}{s_{e'}+1}$$

The instance continuity Γ_i reflects how continuous its component events are. Such that $\Gamma_i(a_i^j)$, for an instance j of a

variation a_i will be defined as in, where $|a_i^j|$ is the length of a_i^j

$$\Gamma_i(a_i^j) = \frac{1}{|a_i^j|} \sum_{k=1}^{a_i^j} \Gamma_e(k)$$

The continuity of a variation, Γ_v, is then defined as the average continuity of its instances. $\Gamma_v(a_i)$ is defined as in, where n_{a_i} shows the total number of instances for variation a_i

$$\Gamma_v(a_i) = \frac{1}{n_{a_i}} \sum_{j=1}^{n_{a_i}} \Gamma_i(a_i^j)$$

The continuity, Γ_g, of a general pattern g is defined as the weighted average continuity of its variations. Γ_g is defined according to, where the continuity for each a_i is weighted by its frequency f_{a_i} and n_a shows the total number of variations for general pattern a

$$\Gamma_g(a_i) = \frac{\sum_{i=1}^{n_a} \Gamma_v(a_i) * f_{a_i}}{\sum_{i=1}^{n_a} f_{a_i}}$$

Building on this definition of continuity, patterns that are interesting and the variation of those patterns are found. The rest of the patterns and variations are pruned.

B. Clustering Sequences into Groups of Activities

The second step of the ADM algorithm is to identify pattern clusters that will represent the set of discovered activities. Specifically, ADM groups the set of discovered patterns P into a set of clusters A. The resulting set of clusters centroids represents the activities that to be modelled, recognize, and track. Though ADM uses a standard k-means clustering method [10], there is a need to define a method for determining cluster centroids and for comparing activities in order to form clusters. Two methods that are commonly used for comparing the similarity of sequences are edit distance and longest common subsequence (LCS) for simple sequences.

The patterns discovered by DVSM were composed of sensor events. In the clustering algorithm, the pattern is composed of states. States correspond to the pattern's events, but are enhanced to include additional information such as the type and duration of the sensor events. In addition, several states together to form a new state. Then all consecutive states are combined that are corresponding to the sensors of the same type to form an extended state. To calculate the similarity between two activities X and Y, compute the distance between their extended state lists and using our general edit distance to account for the state information and the order mapping frequencies. The general edit distance for two patterns X and Y can be defined based on the traditional edit distance, the order distance, and the attribute distance

The general edit distance gives us a measure to compare activities and also to define cluster centroids. Each cluster representative represents a class of similar activities, forming a compact representation of all the activities in the cluster. The activities represented by the final set of clusters are those that are modelled and recognized by the CASAS smart environment. It should be noted that currently, the number of clusters is provided to the clustering algorithm. However, alternative methods can be used to determine the number of clusters during runtime, by forming incremental clusters until no more change can be perceived.

RECOGNIZING ACTIVITIES

Once the activities are discovered, a model has to build for activity recognition. In our approach, Hidden Markov model is used to recognize activities from sensor data. Each model is trained to recognize the patterns that correspond to the cluster representatives found by ADM. A separate Markov model could be learned for each activity and the model that supports a new sequence of events would be selected as the activity label for the sequence.

For this task, hidden Markov model is used, which is a statistical model in which the underlying data are generated by a stochastic process that is not observable. HMMs perform well in the cases where temporal patterns need to be recognized. As with a Markov chain, the conditional probability distribution of any hidden state depends only on the value of a finite number of preceding hidden states.

An HMM model is specified, that using three probability distributions: the distribution over initial states, the state transition probability distribution and the observation distribution. The most likely sequence of hidden states are found that will be given to the observation in and by using the Viterbi algorithm [11].

One drawback of these HMMs sometimes it makes a very slow transition from one activity to another. To remedy this problem, an event-based sliding window is used and this limits the history of sensor events that the model remembers at any given time.

For activity recognition, a voting multi-HMM model is used as a boosting mechanism. Then multiple HMMs is constructed and recognize activities by combining their classifications using a voting mechanism. Specifically, the first HMM represents the first variation of all patterns (one hidden state per pattern), the second HMM represents the second variation of patterns, and so on.

The Viterbi algorithm is used for each HMM to identify the sequence of hidden states, one hidden state at a time, and then, using the described voting mechanism, then identify the most likely hidden state for the multi-HMM based on input

from all individual HMMs. The multi-HMM is built automatically using the output of ADM's discovery and clustering algorithm.

PROPOSED WORK

Fuzzy Sate Q-Learning

To increase the prediction accuracy a Fuzzy-state Q-learning algorithm (FSQL) is proposed that is capable of learning a sequence of actions on the basis of the structure discovered by the process of DVSM. It is an extended version of the Modified Q-Learning Method with Fuzzy State Division, and aims to deal with the states under uncertain conditions, based on discovery of concurrent patterns from data. The procedure starts with some given fuzzy partitions. The number of fuzzy partitions decides the number of linguistic descriptions that are needed to reflect the model's complexity. So that the prediction accuracy is increased.

A. Fuzzyfication

First step in fuzzy logic is to convert the measured data into a set of fuzzy variables. It is done by giving value (these will be our variables) to each of a membership functions set. Membership functions take different shape. Different membership functions are used here is Triangular membership function, Trapezoidal Function, Gaussian membership function.

B. Fuzzy Rules and Inference System

The fuzzy inference system uses fuzzy equivalents of logical AND, OR and NOT operations to build up fuzzy logic rules. An inference engine operates on rules that are structured in an IF-THEN format. The IF part of the rule is called the antecedent, while the THEN part of the rule is called the consequent. Rules are constructed from linguistic variables. These variables take on the fuzzy values or fuzzy terms that are represented as words and modelled as fuzzy subsets of an appropriate domain.

C. Defuzzyfication

The last step of a fuzzy logic system consists in turning the fuzzy variables generated by the fuzzy logic rules into real values again which can then be used to perform some action. There are different defuzzification methods are Centroid Of Area (COA), Bisector Of Area (BOA), Mean Of Maximum (MOM), Smallest Of Maximum (SOM) and Largest Of Maximum (LOM).

The main advantages of using fuzzy logic system are the simplicity of the approach and the capacity of dealing with the complex data acquired from the different sensors. Fuzzy set theory offers a convenient way to do all possible combinations with these sensors. Fuzzy set theory is used in this system to monitor and to recognize the activities of people within the environment in order to timely provide support for safety, comfort, and convenience. Automatic health monitoring is

predominantly composed of location and activity information. Abnormality also could be indicated by the lack of an activity or a abnormal activity detection which will cause or rise the home anxiety.

The first step for developing this approach is the fuzzification of system outputs and inputs obtained from each sensor and subsystem. These membership functions are ordered, firsly according to the area where they maybe occur and secondly according to the degree of similarity between them. The next step of our fuzzy logic approach is the fuzzy inference engine which is formulated by a set of fuzzy IF-THEN rules.
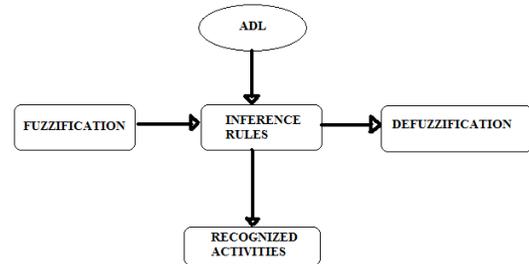
Figure.2 ADL Recognition Method

This second stage uses domain expert knowledge regarding activities to produce a confidence in the occurrence of an activity. Rules allow the recognition of common performances of an activity, as well as the ability to model special cases. A confidence factor is accorded to each rule and in order to aggregate these rules the Mamdani or Sugeno approaches are available under our fuzzy logic component. After rules aggregation the defuzzification is performed by the centroid of area for the ADL output. This framework also allows for rules to be added, deleted, or modified to fit each particular resident based on knowledge about their typical daily activities. This approach based on fuzzy logic provides robust and high accuracy recognition rate on the discovered data [12].

RESULTS

For discovering the activities in a smart environment, Discovering Frequent Discontinuous Activities method (3.1) finds the frequent patterns which are sequentially occurring in the smart environment and it eliminates the variant patterns.
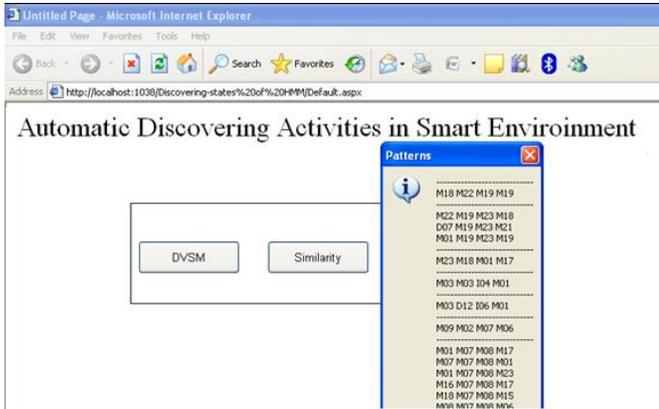
Figure.3 Intersted Pattern of activties

These interesting patterns are given to the Clustering Sequence into Groups of Activities method (3.2) which takes the Sensor ID and the Time as attributes for grouping the similar activities. Similarity is calculated based on mapping between the activities. These clustered activities are given to the Activity Recognition Method (3.3), it uses the HMM model which finds the probability between the activities by forming the states of activities.

Figure.4 Pruned Patterns

Figure.5 Clustered patterns
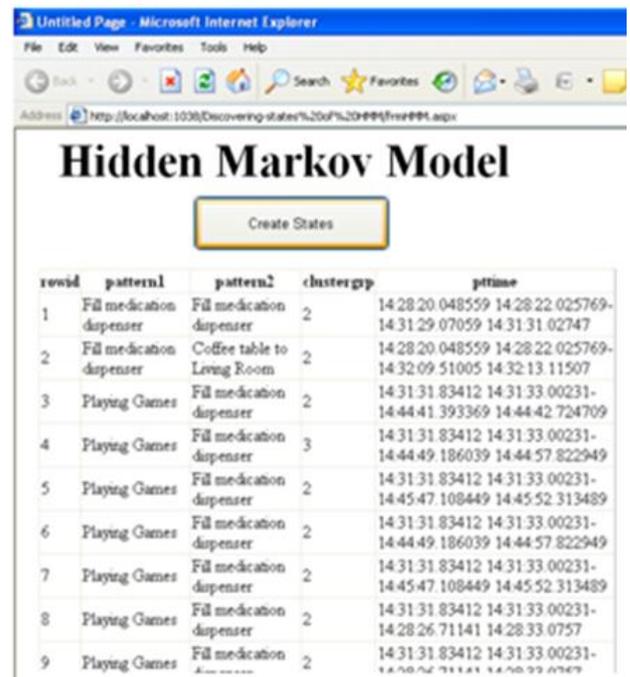
Figure.6 States of Hidden Morkov Model

CONCLUSION

In this paper, ADL activities of smart home residents are tracked and recognized by using the Activity Discovery Method (ADM), which is used to discover the frequent patterns that naturally occur in the smart home. These frequent patterns are clustered for recognizing the activities. For recognizing the activities Multi HMM model is used, which is a supervised

model and it has the difficulty while analyzing the complicated patterns. So, the Fuzzy State Q-Learning method will be used instead of HMM model to produce the effective result for recognizing the activities.

REFERENCE

[1] B.Reisberg, S. Finkel, J. Overall, N. Schmidt-Gollas, S.Kanowski, H. Lehfeld, F. Hulla, S.G. Sclan, H.-U. Wilms, "The Alzheimer's Disease Activities of Daily Living International Scale (ADL-IS)," Int'l Psychogeriatrics, vol. 13, no. 2pp. 163-181, 2001.

[2] V.G.Wadley, O. Okonkwo, M. Crowe, and L.A. Ross-Meadows, "Mild Cognitive Impairment and Everyday Function: Evidence of Reduced Speed in Performing Instrumental Activities of Daily Living," Am. J. Geriatric Psychiatry, vol. 16, no. 5, pp. 416-424, May2008.

[3] E.M.Tapia, S.S. Intille, and K. Larson, "Activity Recognition in the Home Using Simple and Ubiquitous Sensors," Pervasive Computing, vol. 3001, pp. 158-175, 2004.

[4] M.Philipose, K.P. Fishkin, M. Perkowitz, D.J. Patterson, D. Fox, H.Kautz, and D. Hahnel, "Inferring Activities from Interactions with Objects," IEEE Pervasive Computing, vol. 3, no. 4, pp. 50-57, Oct.-Dec. 2004.

[5] O.Brdiczka, P. Reignier, and J.L. Crowley, "Detecting Individual Activities from Video in a Smart Home," Proc. 11th Int'l Conf. Knowledge-Based and Intelligent Information and Eng. Systems (KES), pp. 363-370, 2007.

[6] T.van Kasteren and B. Krose, "Bayesian Activity Recognition in Residence for Elders," Proc. Third IET Int'l Conf. Intelligent Environments (IE '07), pp. 209-212, Sept. 2007.

[7] A.Ruotsalainen and T. Ala-Kleemola, "GAIS: A Method for Detecting Discontinuous Sequential Patterns from Imperfect Data," Proc. Int'l Conf. Data Mining, pp. 530-534, 2007.

[8] V.L.Levenshtein, "Binary Codes Capable of Correct Deletions, Insertions, and Reversals," Soviet Physics Doklady, vol. 10, no. 8, pp. 707-710, 1966.

[9] D.J.Cook and P. Rashidi, "The Resident in the Loop: Adapting the Smart Home to the User," IEEE Trans. Systems, Man, and Cybernetics, Part A: Systems and Humans, vol. 39, no. 5, pp. 949-959, Sept. 2009.

[10] J.A.Hartigan and M.A. Wong, "A K-Means Clustering Algorithm,"Applied Statistics, vol. 28, pp. 100-108, 1979.

[11] A.Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," IEEE Trans. Information Theory, vol. IT-13, no. 2, pp. 260-269, Apr. 1967.

NEURAL NETWORK BASED UNIFIED POWER QUALITY CONDITIONER

N.Ramchandra[1], M.Kalyanchakravarthi[2]

[1] (Student, Department of Electrical and Electronics Engineering, KL University, India)
[2] (Assistant Professor, Department of Electrical and Electronics Engineering, KL University, India)

ABSTRACT

The application of artificial intelligence is growing fast in the area of power sectors. The artificial neural network (ANN) is considered as a new tool to design control circuitry for power-quality (PQ) devices. In this paper, the ANN-based controller is designed and trained offline using data from the conventional proportional-integral controller. The performances of ANN and PI controller are studied and compared for Unified Power Quality Conditioner using MATLAB Simulations.

Key Words: **Artificial intelligence (AI), Artificial neural network (ANN), Current Source Inverter (CSI), proportional integral (PI), unified power-quality conditioner (UPQC), Voltage Source Inverter (VSI)**

I.INTRODUCTION

The use of electronic controllers in the electric power-supply system has become very common. These electronic controllers behave as nonlinear load and cause serious distortion in the distribution system and introduce unwanted harmonics in the supply system, leading to decreased efficiency of the power system network and equipment connected in the network [1]. To meet the requirements of harmonic regulation, passive and active power filters are being used in combination with the conventional converters [2]. Presently, active power filters (APFs) are becoming more affordable due to cost reductions in power semiconductor devices, their auxiliary parts, and integrated digital control circuits. In addition, the APF also acts as a power-conditioning device which provides a cluster of multiple functions, such as harmonic filtering, damping, isolation and termination, load balancing, reactive-power control for power-factor correction and voltage regulation, voltage-flicker reduction, and/or their combinations.

Resent research focuses on use of the universal power quality conditioner (UPQC) to compensate for power-quality problems [3], [4]. The performance of UPQC mainly depends upon how accurately and quickly reference signals are derived. After efficient extraction of the distorted signal, a suitable dc-link current regulator is used to derive the actual reference signals. Various controlling devices like PI, PID, fuzzy logic, and sliding-mode are in use. The basic disadvantage of PI and PID controllers are these need precise linear mathematical models. These fail to operate when non linear conditions are applied. In the recent years Artificial-intelligence (AI) techniques, particularly the NNs, are having a significant impact on power-electronics applications. Neural-network-based controllers provide fast dynamic response while maintaining the stability of the converter system over a wide operating range and are considered as a new tool to design control circuits for PQ devices [5]–[8]. A lot of research works are going on UPQC combined with neural network. In this paper design of ANN based controller is designed for current control and voltage control of shunt active filter instead of PI controller. Two cases are considered where current source inverters and voltage source inverters are taken ANN controller and PI controller performances at the DC link are compared using MATLAB/SIMULINK.

II.UNIFIED POWER QUALITY CONDITIONER

A conventional UPQC topology consists of the integration of two active power filters are connected back to back to a common dc-link bus [9]. A simple block diagram of a typical UPQC is shown in Fig. 1.
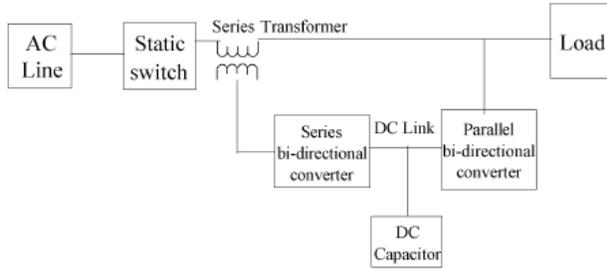
Fig 1: Block diagram of UPQC

UPQC with current source inverters and voltage source inverters are shown in fig 2 and fig 3 respectively
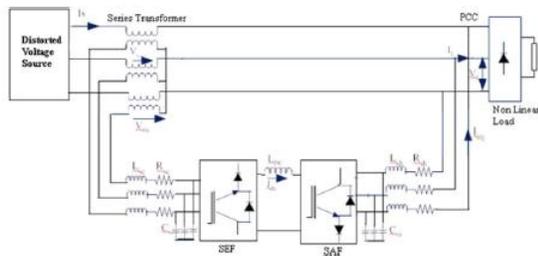
Fig 2: UPQC topology with current source inverters
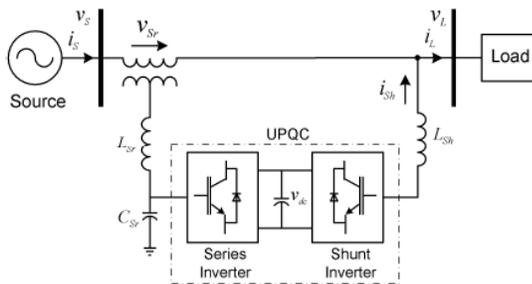
Fig 3: UPQC topology with voltage source inverters

It can be configured either with voltage-source converters or current source converters in single phase, three-phase three wire, or three-phase four-wire configurations. The UPQC with the voltage-source converter (VSC) is most common because of its smaller size and low cost. The current source inverters are used because of their excellent current control capability, easy protection, and high reliability. The performance of the UPQC mainly depends on how accurately and quickly the reference signals are derived [11]. The DC link will derive the reference signals. The difference between actual and reference signals are given to controller whether it

may be PI or any other controller. The output of controller is given to the generation of pulses.

III.DESIGN OF THE PI CONTROLLER

The design of PI controller for current source inverter is done by the following assumptions
1) The voltage at PCC is sinusoidal and balanced.

2) Since the harmonic component does not affect the average power balance expressions, only the fundamental component of currents is considered.

3) Losses of the system are lumped and represented by an equivalent resistance connected in series with the filter inductor.

4) Ripples in the dc-link current are neglected.
The average rate at which energy being absorbed by the inductor is

$$Pind = \frac{d}{dt}\left(\frac{1}{2}LdcIdc^2\right) = LdcIdc\frac{dIdc}{dt} \qquad (1)$$

The power input to the PWM converter

$$Pconv = 3VshIinj \qquad (2)$$

The average rate of change of energy associated with the capacitor filter

$$Pcap = \frac{d}{dt}\left(\frac{1}{2}CshVsh^2\right) \qquad (3)$$

Power loss in the resistor Rsh

$$Ploss = 3Iinj^2Rsh \qquad (4)$$

Equating them

$$Pind = Pconv - Ploss - Pcap \qquad (5)$$

Substituting the values

$$LdcIdc\frac{dIdc}{dt} = 3(VshIinj - Iinj^2Rsh - CshVsh\frac{dVsh}{dt})$$

$$(6)$$

In order to linearize the power equation, a small perturbation ΔIinj is applied in the input current Iinj of converter about a steady-state operating point Iinjo, the average dc-link current will also get perturbed by a small amount ΔIdc around its steady-state operating point Idco

$$Iinj = Iinjo + \Delta Iinj \quad \text{and} \quad Idc = Idco + \Delta Idc$$

In (6) neglecting higher order terms

$$LdcIdco\frac{d\Delta Idc}{dt} = 3(VshIinjo + Vsh\Delta Iinj - Iinjo^2 Rsh - 2Iinj\Delta IinjoRsh - CshVsh\frac{dVsh}{dt})$$
(7)

Subtracting (7) from (6)

$$LdcIdc\frac{d\Delta Idc}{dt} = 3(Vsh\Delta Iinj - 2Iinj\Delta IinjoRsh - CshVsh\frac{dVsh}{dt}$$
)
(8)

The transfer function of the PWM converter for a particular operating point

$$Kc = \frac{\Delta Idc}{\Delta Iinj} = 3(\frac{Vsh - CshVshS - 2IinjoRsh}{LdcIdcoS})$$
(9)

The characteristic equation of Pi controller is

$$1 + \left(Kp + \frac{Ki}{S}\right)3\left(\frac{Vsh - CshVshS - 2IinjoRsh}{LdcIdcoS}\right) = 0$$

Some of the parameters are taken from [11] as

Vsh=230V, Iinjo=5 amp, Rsh=0.4Ω, Csh=24μF, Ldc=160mH, Idco=5 amp

Hence the characteristic equation on substitution of the values is

$$0.8S^2 + kp(678 - 0.0165S^2) + Ki(678 - 0.0165S)$$
(10)

Using Routh Harwitz criteria the values of Kp=0.5 and Ki=10 are chosen for the PI controller that is used in UPQC.

IV. DESIGN OF ANN CONTROLLER

The rapid detection of the disturbance signal with high accuracy, fast processing of the reference signal, and high dynamic response of the controller are the prime requirements for desired compensation in case of UPQC. The conventional controller fails to perform satisfactorily under parameter variations nonlinearity load disturbance, etc. A recent study shows that NN-based controllers provide fast dynamic response while maintaining stability of the converter system over wide operating range.

The ANN is made up of interconnecting artificial neurons. It is essentially a cluster of suitably interconnected nonlinear elements of very simple form that possess the ability to learn and adapt. It resembles the brain in two aspects: 1) the knowledge is acquired by the network through the learning process and 2) interneuron connection strengths are used to store the knowledge [10]-[11]. These networks are characterized by their topology, the way in which they communicate with their environment, the manner in which they are trained, and their ability to process information. ANNs are being used to solve AI problems without necessarily creating a model of a real dynamic system. For improving the performance of a UPQC, a multilayer feed forward-type ANN-based controller is designed. This network is designed with three layers, the input layer with 2, the hidden layer with 21, and the output layer with 1 neuron, respectively.

The training algorithm used is Levenberg–Marquardt back propagation (LMBP). The MATLAB programming of ANN training is given as follows:

```
net=newff(minmax(P),[2,21,1],{'tansig','tansig','purel
in'},'trainlm');
net.trainParam.show=50;
net.trainParam.lr=0.05;
net.trainParam.mc=0.95;
net.trainParam.lr_inc=1.9;
net.trainParam.lr_dec=0.15;
net.trainParam.epochs=5000;
net.trainParam.goal=1e-6;
[net,tr]=train (net,P,T);
a=sim (net,P);
gensim (net,-1);
```

V.SIMULATION RESULTS
1) UPQC with current source inverters

The system considered is 3-phase system and load is taken as non linear load. UPQC consists of series inverter and shunt inverter which are current source inverters. An inductor is taken as taken as a dc link between the inverters.

The parameters of transmission line are taken same values which are mentioned design of PI controller. The simulation diagram is shown in figure 4.
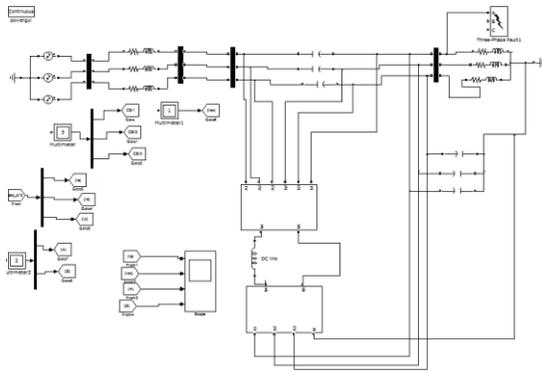

Fig 4: Simulation diagram UPQC with Current Source Inverters

The simulation is performed for 0.3sec and a disturbance at load is applied for a certain period of time and the performance of PI and ANN controller are compared.
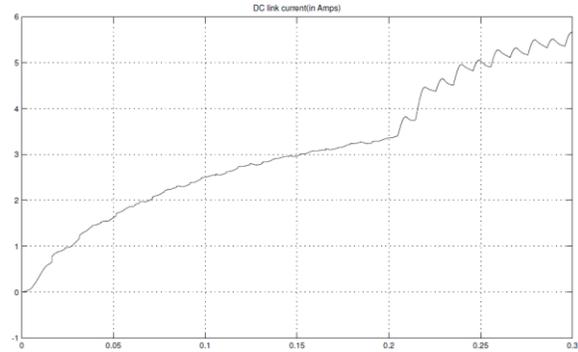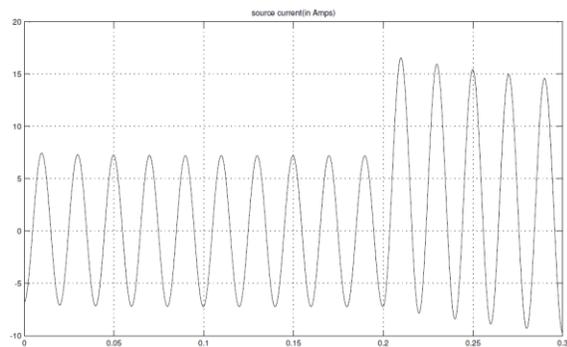
Fig 5: Performance of UPQC with PI controller at load perturbations
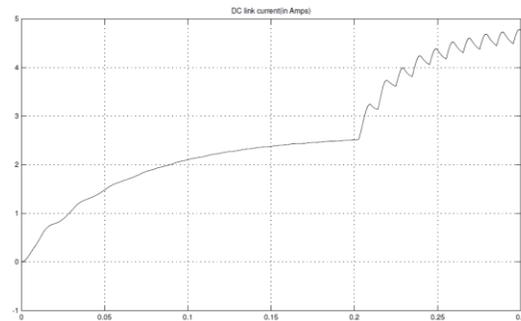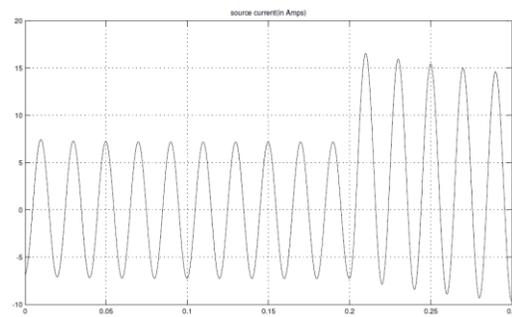
Fig 6: Performance of UPQC with ANN controller at load perturbations

Total harmonic distortion is also taken (0.15sec and 0.25sec). PI and ANN controller performance is compared
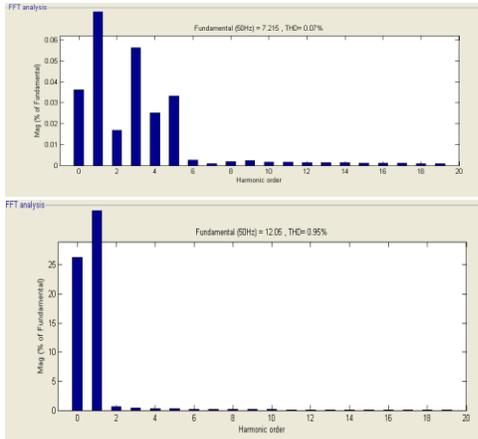
Fig 7: Frequency spectrum of the source current at different loading conditions with the PI controller.
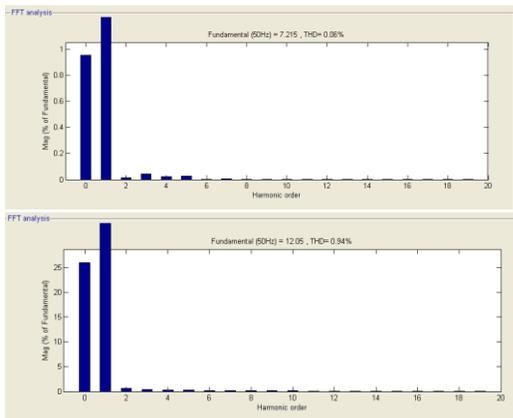
Fig 8: Frequency spectrum of the source current at different loading conditions with the ANN controller.

From figures 5 and 6 the dc link current is taking more to stabilize at initial conditions and load perturbations in the case of PI controller and in the other case of ANN controller dc link current is stabilizing fast in both conditions compared to PI controller.

The performance of harmonic current filtration is shown. The load current in both cases is found to be content of all odd harmonic minus triplen, providing a total harmonic distortion (THD) of 27.82%. It is observed from the figure that the THD of the source current at 0.15 s is 0.07% in the case of the PI controller while it is 0.06% in the case of the ANN controller scheme. Similarly, the THD of the source

current at 0.25 s is 0.95% in case of the PI controller while it is 0.94% in case of the ANN controller scheme. At both cases ANN controller performance is proving better than PI controller.

2) UPQC with voltage source inverters

The system considered is 3-phase system and load is taken as non linear load. UPQC consists of series inverter and shunt inverter which are voltage source inverters. A capacitor is taken as taken as a dc link between the inverters.
The parameters of transmission line are taken same values which are mentioned design of PI controller. The simulation diagram is shown in figure 9.
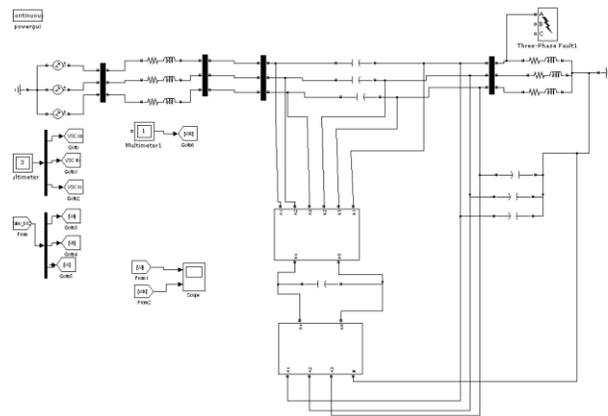
Fig 9: Simulation diagram UPQC with Voltage Source Inverters

The simulation is performed for 0.3sec and a disturbance at load is applied for a certain period of time and the performance of PI and ANN controller are compared
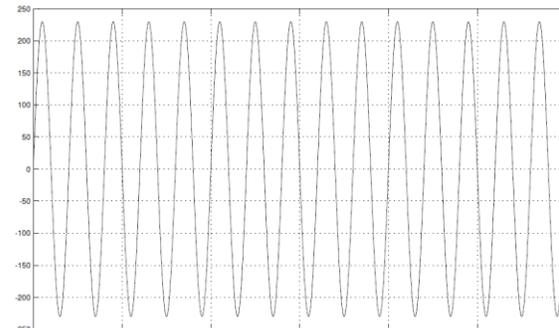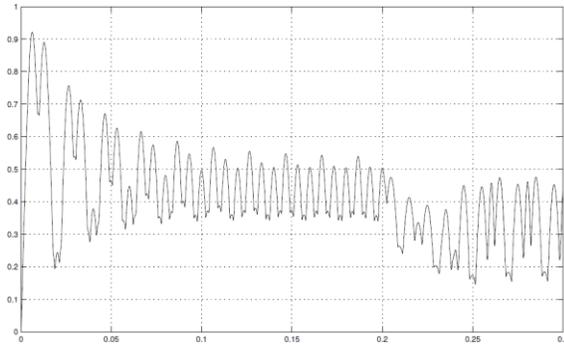
Fig 10: Performance of UPQC with PI controller at load perturbations
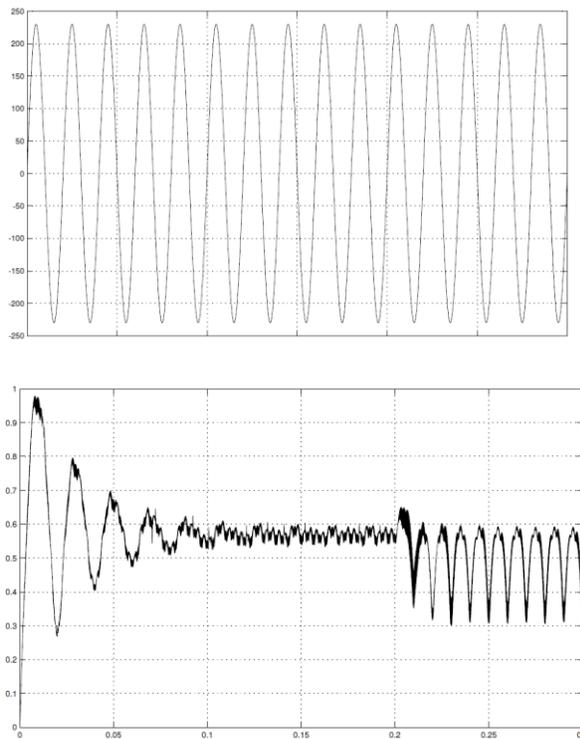
Fig 11: Performance of UPQC with ANN controller at load perturbations

From figures 10 and 11 the dc link is stabilizing fast at initial conditions with ANN controller compared to PI controller. Even at load perturbations there are fewer oscillations occurring with ANN controller compared to PI controller. Hence ANN controller is showing a better performance in the two cases against PI controller.

VI.CONCLUSION

The performance of the UPQC mainly depends upon how accurately and quickly reference signals are derived. There were several conditions that are tested. However, the performance of conventional PI controller is not proving better against proposed ANN controller in both cases of UPQC (considering CSI and VSI inverters). This is proved through simulation results. Finally, with ANN controller there was considerable improvement in the response time of the control of the dc-link current which is the main issue in the case of the power system network.

REFERENCES

[1] E. W. Gunther and H. Mehta, "A survey of distribution system power quality," *IEEE Trans. Power Del.*, vol. 10, no. 1, pp. 322–329, Jan. 1995

[2] W. M. Grady, M. J. Samotyj, and A. A. Noyola, "Survey of active power line conditioning methodologies," *IEEE Trans. Power Del.*, vol. 5, no. 3, pp. 1536–1542, Jul. 1990.

[3] F. Kamron, "Combined dead beat control of series—Parallel converter combination used as a universal power filter," in *Proc. IEEE Power Electronics Specialist Conf.*, 1995, pp. 196–201.

[4] H. Fujita and H. Akagi, "The unified power quality conditioner: The integration of series active filter and shunt active filters," in *Proc. IEEE/ Power Eng. Soc. Power Electronics Specialist Conf.*, Jun. 1996, pp. 491–501

[5] A. Zouidi, F. Fnaiech, and K. AL-Haddad, "Neural network controlled three-phase three-wire shunt active power filter," in *Proc. IEEE ISIE*, Montreal, QC, Canada, Jul. 9–12, 2006, pp. 5–10.

[6] R. El Shatshat, M. M. A. Salama, and M. Kazerani, "Artificial intelligent controller for current source converter-based modular active power filters," *IEEE Trans. Power Del.*, vol. 19, no. 3, pp. 1314–1320, Jul. 2004.

[7] J. R. Vazquez and P. R. Salmer6n, "Three-phase active power filter control using neural networks," in *Proc. 10th Mediterranean Electro Technical Conf.*, 2000, vol. III, pp. 924–927.

[8]   A. Elmitwally, S. Abdelkader, and M. EL-Kateb, "Neural network controlled three-phase four-wire shunt active power filter," *Proc. Inst. Elect. Eng., Gen. Transm. Distrib.*, vol. 147, no. 2, Mar. 2000.

[9]   A. Nasiri and A. Emadi, "Different topologies for single-phase unified power quality conditioner," in *Proc. Conf. Rec. Industry Applications*, 2003, pp. 976–981.

[10] K. Sunat, *Neural Networks and Theory and Applications*, ser. Lecture Notes. India: Burapha Univ., Jul. 2, 2006

[11] Vadirajacharya G. Kinhal, Promod Agarwal, and Hari Oam Gupta, "Performance Investigation of Neural-Network-Based Unified Power-Quality Conditioner" IEEE Trans. ON POWER DELIVERY, VOL. 26, NO. 1, JANUARY 2011

Improving E-learning System using Ontology Web Language

Priya.L[1], Ravikumar.G[2], Anand Kumar.M[3], Dr.Gunasekaran.S[4]
Kanimozhi.E[5], Jennifer Diana.C[6]

[1,5,6]*Post Graduate Student*
[2,3]*Assistant Professor*
[4]*Head of the Department*
Department of CSE
Coimbatore Institute of Engineering and Technology
Coimbatore, Tamilnadu, India

Abstract— **In recent years, most of the conventional education are suitable for requirements of progress in educational development but they don't able to cope up with the changes of learning demand in time, thus computer networks brought an opportunity for it. E-learning is one of the best solutions for it and is used to represent a wide spectrum of application, ranged from virtual class rooms to remote course or distance learning. Web based courses helps the learners by making access to the educational resources very fast, just-in-time and relevance, which is not depending on time and place. Previously semantic web based model for e-learning system uses XML but nowadays it is developed with RDF data model and OWL ontology language. By combining RDF and OWL into e-learning, learning become feasible and this technology can greatly improve the efficiency of learning and achieve a win-win situation between instructor and learners.**

Keywords—   **E-learning, SW-Semantic Web, Ontology , RDF-Resource Description Framework, OWL- Web Ontology Language, , DAML-DARPA Agent Mark-up Language, OIL-Ontology Interface Language**

I.   INTRODUCTION

E-learning is an Internet-based learning process, it uses internet technology to design, implement, select, manage, support and extend learning, which will not replace traditional education methods, but will greatly improve the efficiency of the education. E-learning has a lot of advantage like flexibility, diversity, measurement, openness and etc. it will become a primary way for learning in new century.

Artificial Intelligent (AI) along with internet technology is known as semantic web which is the most interesting and evolving technology for e-learning. It is about making the web more understandable by the machines through an appropriate infrastructure for the intelligent agents to move around the web to perform any complex task for their user.

A semantic web is a process of creating the web as machine-understandable and interoperable service that an intelligent agent of AI can be discovered, executed, and composed automatically. Obviously the web was build for human consumption and not for machine consumption, (i.e.) it is machine-readable and not machine –understandable. The semantic web based application enable to interoperate both on semantic level and syntactic level, this will help the semantic web to express information in a precise, machine readable form and help the software agents to process, share and reuse it.

II.   SEMANTIC WEB

It is a ―Web-of-data that facilitate the machine to understand the semantics or meaning of information on the World Wide Web. It extends the network of hyperlinked human readable web pages by inserting machine readable metadata and how these data are related to each other. There are four categories of important issues that related to semantic web,

1. SW languages.
2. Ontologies
3. Semantic mark-up of web pages
4. Semantic web services

A.   Layers of The Semantic Web

The main goal of semantic web is to express the meaning of the content. In order to achieve the goal several layers are needed indeed. They are represented in Figure 1.
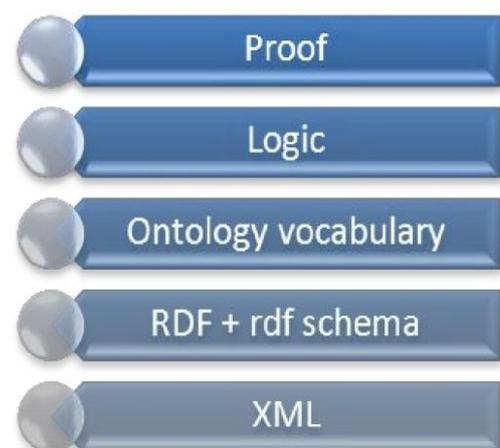
Fig 1 Layers of the Semantic web architecture

The XML layer, which represent the structure of data, the RDF layer, represent the meaning of data, the Ontology layer, represent the formal common agreement about meaning of data, the Logical layer, enables intelligent reasoning with meaningful data

B. Semantic Web Languages

A specific programming language is necessary for representing information on semantic web and also makes that information useful in syntactically and semantically. The languages used in modern semantic web technology are XML (eXtensible Mark-up Language), XML schemas, RDF (Resource Description framework) and RDF schemas. All are developed by w3c by using XML syntax.

An XML document can be viewed as a tree, where leaf corresponds to data values and nodes represents XML elements. It helps user to create its own tags. An XML document consists of three parts; first one is an XML declaration, second is a DTD or XML schema and third is an XML instance (XML document data). The XML declaration specifies the version and the encoding of XML being used, and Document type definition (DTD) helps to structure the content of the documents.

The Resource Description Framework (RDF) is a standard model for interchanging data on the web. RDF has a significant feature that facilitate data merging even if the underlying schemas are differ for adding semantics to a document. It is an infrastructure that enables encoding, exchange and reuse of structured metadata. In general, information is stored in the form of RDF statements, which are easily understandable by search engines, intelligent agents, browsers and human beings to use the semantic information.

C. Ontologies

Information on the web is commonly represented in Natural Languages for human understanding, but not for the computers. It is necessary to represent a language in a form that can be interpreted syntactically and semantically by a computer. Ontology is one of the best key for providing information in a computer-understandable way. It is defined as a formal, explicit representation of the objects and relation and specification of a shared conceptualization. It also defines as the common vocabulary for the researchers to share information in a domain. Ontologies are applied to web for creating the semantic web. It typically consists of definitions which are relevant to the domain, their relations, and axioms about the concepts and relationships.

OWL (web ontology language) is a language that was released by w3c for representing ontology. It is developed from description logic and DAML + OIL and they are developed using integrated, graphical, ontology- authoring tools where DAML and OIL are DARPA Agent Mark-up Language, Ontology Interface Language respectively. DAML+OIL is a successor language to DAML and OIL that combines these features from both. In turn, it was superseded by Web Ontology Language (OWL). The DAML + OIL language has also been developed as an extension of XML and RDF. Ontology as a formal semantic account, see Figure

2, which is analysed the phenomenon of e-Learning and have concluded several semantic that formulate a value layer capable of exploiting knowledge sources semantically. The major problem concerning this interpretation of ontology is the complexity of e-Learning.

DAML stands for DARPA Agent Mark-up Language. OIL stands for Ontology Inference Layer or Ontology Interchange Language. The increasing popularity of OWL might lead to its widest adoption as the standard ontology representation language on the semantic web. The problem of complexity of e-learning gets reduced by DAML.

D. Semantic Mark Up

In order to make the web content to machine understandable, web pages and the documents must contain semantic mark up. For performing this annotation Knowledge Annotator Tool is used which is a standalone tool available in the internet world.
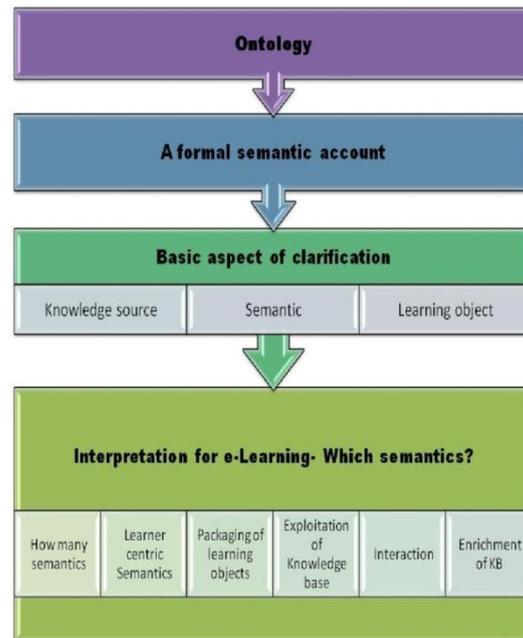
Fig 2 Ontology as a formal semantic account

E. Semantic Web Services

There are few features that the user wants from the semantic web; they must be intelligent and provides high level services like information brokers, search agents, information filters, intelligent information integration and knowledge management. These features are possible only when ontologies get populated on the web which will help the interoperation between the agents and the application on the semantic web.

II. SEMANTIC WEB AND E-LEARNING

The great success of the current WWW leads to a new challenge that a huge amount of data is interpretable by humans only with limited machine support.

TABLE I
COMPARISON OF E-LEARNING AND SEMANTIC WEB WITH FEATURES

Requirements	e-learning	Semantic web
Delivery	Pull- Student determines Agenda	Knowledge Items (learning materials) are distributed on The web, but they Are linked to Commonly agreed ontologie(s). This Enables construction of a user-specific course, By semantic querying For topic of interest.
Responsiveness	Reactionary- Responds to problem at hand	Software agents on the Semantic Web May use a commonly agreed Service language, Which enables co-ordination between agents And proactive delivery Of learning materials In the context of Actual problems. The Vision is that each user has his own Personalized agent That Communicates With other agents.
Access	Non-linear – Allows direct access to knowledge in whatever sequence makes sense to the situation at hand	User can describe the situation at hand (goal of learning, Previous knowledge,...) And perform Semantic querying for The suitable Learning material. The User profile Is Also accounted for. Access to knowledge can Be expanded by Semantically Defined navigation.
Symmetry	Symmetric – Learning occurs as an integrated activity	The Semantic Web (semantic intranet) offers the potential to become an integration platform for All business processes in an organization, including Learning activities.
Modality	Continuous –	Active delivery Of

	Learning runs in parallel to business tasks and never stops	information (based on personalized agents) creates a dynamic learning environment that is integrated in the business processes.
Authority	Distributed – Content comes from the interaction of the participants and the educators	The Semantic Web will be as decentralized as possible. This enables an effective co-operative content management.
Personalization	Personalized – Content is determined by the individual user's needs and aims to satisfy the needs of every user	A user (using its personalized agent) searches for learning material customized for her/his needs. The ontology is the link between user needs and characteristics of the learning material.
Adaptively	Dynamic – Content changes constantly through user input, experiences, new practices, business rules and heuristics	The Semantic Web enables the use of distributed knowledge provided in various form. Distributed nature of the Semantic Web enables continuous improvement of learning materials.

IV. SEMANTIC WEB IN EDUCATION

Web based education has become a very important branch of education technology and plays a vital role. Nowadays, web-based systems are facing challenges like extensibility, interoperability, the use of domain ontologies, contextualization and consistence of metadata, dynamic sequencing of learning and contents, integration and reuse of content, distribution of services, new models of learning, and so on. Such challenges are related to the attempt to represent the information on the Web in a way that computers can understand and manipulate it. The main goal of the semantic web-based educational system is to use resources available on the Web through standards based technologies in order to accomplish AAAL: Anytime, Anywhere, Anybody Learning.

According to Anderson and Whitelock [10], the Educational Semantic Web is based on three fundamental affordances.

- The capacity for effective information storage and retrieval.
- The capacity for nonhuman autonomous agents to augment the learning and information retrieval of human beings.
- The capacity of the Internet to support, extend and expand communication capabilities of humans in multiple formats across the bounds of space and time.
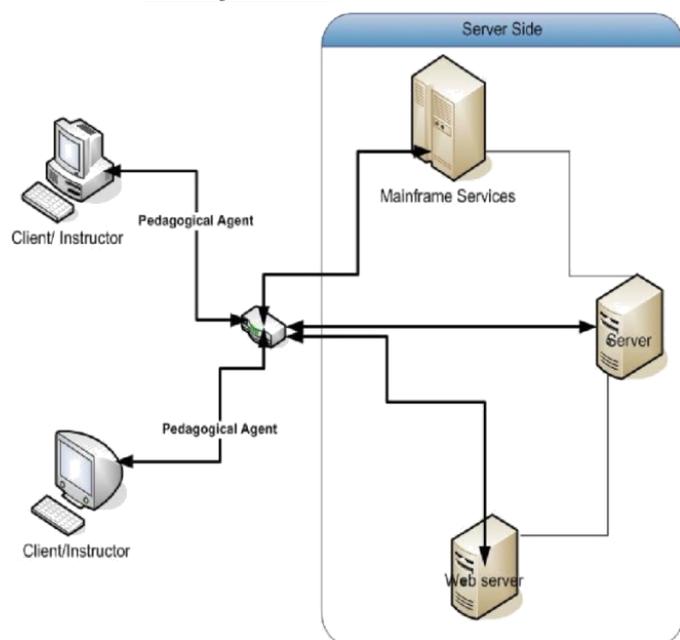
Fig 3 The setting for Semantic Web-based education

The above figure depicts the teaching, learning, collaboration, assessment and other educational activities that are happening on the web. The intelligent pedagogical agents in the Figure 3 help in the flow of information and knowledge between client and servers. These agents are also very much helpful in locating, browsing, selecting, arranging, integrating various education material from different educational servers. These agents access the content from various servers by using high-level educational services as shown in Figure 1.

V. CONCLUSIONS

The biggest growth in the internet, and the area that will prove to be one of the biggest agents of change, will be in E-Learning. E-Learning enables to update materials and information across the entire enterprise, keeping content fresh and relevant. Online training also creates a personalized learning experience. Instead of daylong or weeklong programmers, the typical E-Learning course can be broken into one-hour modules, offering flexibility around training. Employees can adapt training to their own lives and learning styles, accessing material whenever it is convenient to review course material.

REFERENCE

[1] Fayed Ghaleb, Sameh Daoud, Ahmad Hasna, Jihad Jaam, A.EL-Seoud, and Hosam El-Sofany. ―E-Learning model based on semantic web technology, International Journal of computing and information science, vol.4, No.2, August 2006.
[2] Berners-Lee T. ―What the semantic when can represent ,http://www.w3.org/DesignIssues/RDF/not.html, Accessed 10 Aug 2010.
[3] Zuhoor Alkhanjari, Swamy kutti, and Muna Hatem. ―An Extended E-Learning system Architecture: Integrating software tools within the E-Learning Portal, in The International Arab Journal of Information Technology, vol. 3, Jan 2006.
[4] Vladan Devedic. ― Education and the semantic web International Journal of Artificial Intelligent in Education 14, 2004.
[5] Ig Ibert Bittencourt, Seiji Isotani, Evandro Costa, and Riichiro Mizaguchi. ―Research Direction on semantic web and education ,Sciential Interdisciplinary studies in Computer science vol.19, No.1, 59-66, Jan 2008.
[6] Pilar Sancho, Ivan Martinez,and Baltasar Fernandez Manjon.― Semantic web technology applied to e-learning Personalization, Journal of universal computer science, vol.11, No.9, 1470-1481, 2005.
S.Muthulakshmi, G.V.Uma. ―Semantic web-based e-learning system for sports domain, International Journal of Computer Application (0975-8887) vol.8, No.14, October 2010.
Sarah Gutierrez. ―The impact of Semantic Web on education, INF 385T- Semantic Web Technology, 9 Dec 2008.
Horrocks, I. and Van Harmelen, F―Description of the DAML+ OIL Ontology Markup Language, March 19, 2002.
Anderson T. and Whitelock D. ―The educational semantic web: visioning and practicing the future of education, Journal of Interactive Media in Education (JIME), vol.1, 2004.

ABOUT AUTHORS

Ms.Priya. L is currently pursing M.E, CSE in Coimbatore Institute of Engineering and Technology (C.I.E.T), Coimbatore and received her B.E degree in Computer Science and Engineering at Coimbatore Institute of Engineering and Technology under Anna University, Chennai. Her research area is Unsupervised Cross-Lingual Lexical Substitution from Natural Language.

Mr.G.Ravikumar received his M.Tech., degree and B.E., degree in Computer Science and Engineering from Bharathidasan and Sastra University respectively. He is currently working as assistant professor in CSE department at C.I.E.T, Coimbatore. His research areas are Disk Optimization and Compiler Design.

Mr. AnandKumar. M is working as an Assistant Professor in C.I.E.T, Coimbatore. He is pursuing his Ph.D. under Amirta University Coimbatore. He received his M.Tech., degree at Amirta University. His research area is Machine Learning, Computational Linguistics and Machine Translation.

Dr.Gunasekaran.S received his M.E., degree and B.E., degree in Computer Science and Engineering from Anna University and Bharathiyar respectively. He received his Ph.D. degree under ANNA University, Coimbatore. He is currently working as HOD, in CSE department at C.I.E.T, Coimbatore. His research interests accumulate in the areas of Adhoc Networks, Data Mining, Semantic Web Services and Cloud

Systems. He is also interest in modern pedagogies in engineering education.

Ms Kanimozhi.E is currently pursing M.E in CIET, Coimbatore and received her B.E degree in cse at Mepco Schlenk Engineering College under Anna University, Chennai. Her research areas are Web Services, Information Retrieval from Natural Language.

Ms Jennifer Diana.C is currently pursing M.E ,CSE in CIET, Coimbatore and received her B.E degree in CSE at OCET under Anna University, Chennai. Her research areas are web services and Mobile cloud computing.

CLUSTERING IN ADHOC NETWORKS BASED ON LOAD BALANCING FOR DELAY-TOLERANT APPLICATIONS

Bensi Vijitha.J*, Mrs.Sudha.R**

*II M.E CS, Sri shakthi Institute Of Engineering and Technology,Anna University,Coimbatore
**Asst.prof ECE, Sri shakthi Institute Of Engineering and Technology,Anna University,Coimbatore

Abstract:

Delay-tolerant networking (DTN) is an approach to computer network architecture that seeks to address the technical issues in heterogeneous networks that may lack continuous network connectivity. Due to the lack of continuous communications among mobile nodes and possible errors in the estimation of nodal contact probability, convergence and stability become major challenges in distributed clustering in DTMN. Clustering significantly reduces the energy consumption of a cluster. In this paper, a cluster based routing protocol for Delay-Tolerant Mobile Networks (DTMNs) is used. Exponentially weighted moving average (EWMA) scheme is employed for on-line updating nodal contact probability, with its mean proven to converge to the true contact probability. Based on nodal contact probabilities, a set of functions including sync ( ),leave ( ), and join ( ) are devised for cluster formation and gateway selection. The gateway nodes exchange network information and perform routing. It uses clustering's structure to decrease overhead, average end-to-end delay and improve the average packet delivery ratio.

Keywords: **Delay Tolerant Mobile Network (DTMN), Exponential weighted moving average (EWMA).**

1. INTRODUCTION:

Delay tolerant Mobile Network (DTMN), fundamentally opportunistic communication system, where communication links only exist temporally rendering. It is impossible to establish end to end connection for data delivery. Delay-tolerant networking (DTN) is an approach to computer network architecture that seeks to address the technical issues in heterogeneous networks that may lack continuous network connectivity. The Delay-Tolerant Network (DTN) is an occasionally [2] connected network that may suffer from frequent partitions and that may be composed of more than one divergent set of protocol families. Examples of such networks are those operating in mobile or extreme terrestrial environments, or planned networks in space. DTNs span very challenging application scenarios where nodes (e.g., people and wild animals) move around in environments where infrastructures cannot be installed (e.g., military grounds and protected environments).

Some solutions to routing have been presented also for these cases, starting from the basic epidemic routing, where messages are blindly stored and forwarded to all neighboring nodes, generating a flood of messages. Existing routing protocols (AODV)]9\ are must take to a "store and forward" approach, where data is incrementally moved and stored throughout the network in hopes that it will eventually reach its destination. A common technique used to maximize the probability of a message being successfully transferred is to replicate many copies of the message in the hope that one will succeed in reaching its destination. This is feasible only on networks with large amounts of local storage and internodes bandwidth relative to the expected traffic. The drawbacks that are encountered in Delay-Tolerant Mobile Networking are [5]

1. Lack of Connectivity.

2. Lack of Instantaneous End-To-End Paths

3. Very High Number of Messages that are needed to obtain a successful delivery to the right recipient.

2. PROBLEM DEFINITION:

Clustering in DTMN is unique, because the network is not fully connected. Due to the Lack of continuous communication among mobile nodes and possible errors in the estimation of nodal contact probability, convergence and stability becomes a major problem in the Delay-Tolerant Mobile Network (DTMN) [1]. In non clustered Delay –Tolerant Mobile network any node in the network may not able to get contact with the other neighboring node this is because of the nodes will not be having a correct updating in their nodal contact probability and the gateway information therefore they lack communication. As a result the nodes cannot provide end to end delivery of the information. At the same time nodes break up their

communication if they move out of the coverage area.

3. CLUSTERING:

Clustering is a process that divides the network into interconnected substructures [9], called clusters. Each cluster has a cluster head (CH) as coordinator within the substructure. Each CH acts as a temporary base station within its zone or cluster and communicates with other CHs.

In our protocol, there are three possible states for the node: NORMAL, CLUSTERHEAD and GATEWAY. Initially all nodes are in the state of isolated. All nodes maintain the NEIGHBOR table wherein the information about the other neighbor nodes is stored CHs [9].
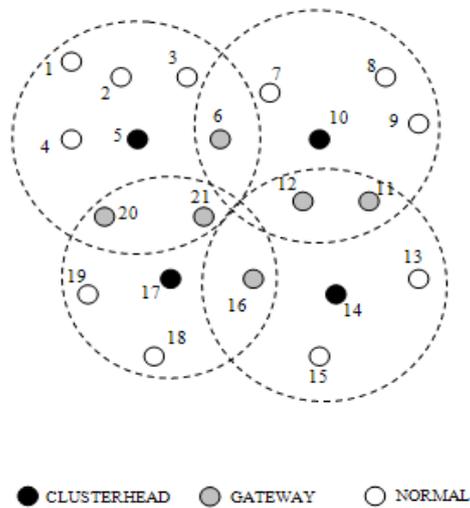
Fig 1: cluster formation

4. DISTRIBUTED CLUSTERING ALGORITHM:

Distributed clustering algorithm used to form a cluster in delay tolerant mobile network. The algorithm is event-driven, where the key part lies on the meeting event between any pair of nodes[6]. The set of functions in the algorithm including Sync, Leave, and Join is outlined below. The Methodology will give out an idea how the algorithm performs its function

METHODOLOGY:

a. In our protocol, an exponentially weighted moving average (EWMA) scheme is employed for on-line updating nodal contact probability.

b. Weighting factors which decrease exponentially. The weighting for each older data point decreases exponentially, giving much more importance to recent observations while still not discarding older observations entirely.

c. True contact probability. Subsequently, a set of functions including *Sync ()*, *Leave ()*, and *Join ()* are devised to form clusters and select gateway nodes based on nodal contact probabilities.

d. Cluster table consists of four fields: Node ID, Contact Probability, Cluster ID, and Time Stamp.

e. Each entry in the table is inserted/updated upon meeting with another node, by using the aforementioned online updating scheme.

f. Gateway table consists of four fields: Cluster ID, Gateway, Contact Probability, and Time Stamp.

NODAL CONTACT PROBABILITY:

The delivery probability indicates the likelihood that r can deliver data messages to the sink. The delivery probability of a power i, is updated as follows,

$$\xi_i = \begin{cases} (1-\alpha)[\xi_i] + \alpha\xi_k, & Transmission \\ (1-\alpha)[\xi_i], & Timeout, \end{cases}$$

where is the delivery probability of power i before it is updated, is the delivery probability of node k (a neighbor of node i), and is a constant employed to keep partial memory of historic status.

A.SYNC

The *Sync ()* process is invoked when two cluster members meet and both pass the membership check. It is designed to exchange and synchronize two local tables. The synchronization process is necessary because each node separately learns network parameters, which may differ from nodes to nodes. The Time Stamp field is used for the "better" knowledge of the network to deal with any conflict.

B.LEAVE

The node with lower stability must leave the cluster. The stability of a node is defined to be its minimum contact probability with cluster members. It indicates the likelihood that the node will be excluded from the cluster due to low contact probability. The leaving node then empties its gateway table and reset its Cluster ID.

C.JOIN

The *Join ()* procedure is employed for a node to join a "better" cluster or to merge two separate clusters. A node will join the other's cluster if, it passes membership check of all current members. Its stability is going to be improved with the new cluster. By joining new cluster, it will copy the gateway table from the other node and update its cluster ID accordingly. Thus the distributed clustering algorithm is used to form a cluster in DTMN.

CLUSTER BASED ROUTING:

The cluster based routing protocol used to perform a routing in delay tolerant mobile network. We consider the Node i has a data message to Node j, the cluster based routing protocol is given below.

A. INTRA-CLUSTER ROUTING

If Nodes i and j are in the same cluster, they have high chance to meet each other, thus Node i will transmit the data message to Node j directly upon their meeting. No relay node is necessarily involved.

B. ONE-HOP INTER-CLUSTER ROUTING

If they are not in the same cluster, Node i look up gateway information to Node j's cluster in its gateway table. If an entry is found, Node i send the data message to that gateway. Upon receiving the data message, the gateway will forward it to *any* node.e.g. Node *k,* in node j's clusters. Node *k,* which in sum delivers the data message to node j via Intra-cluster Routing. If no gateway entry is found, node i precede the Multi-hop Inter-cluster Routing as to be discussed next.

C. MULTI-HOP INTRA CLUSTER ROUTING

If node i does not have any information about node j, the data transmission needs a multi-cluster routing scheme. Given the low connectivity environment, on-demand routing protocols, with extremely high packet dropping probability, will not work effectively here. However, any table-driven routing algorithm such as the following link-state-like protocol can be employed. In the protocol, every gateway node builds a *Cluster Connectivity Packet (CCP),* and distributes it to other gateways in the network.

1. The CCP of a Gateway comprises its cluster ID and a list of clusters to which it serves as gateway along with corresponding contact probabilities. Such information can be readily obtained from the gateway table.

2. Once a gateway node accumulates a sufficient set of CCP's, it constructs a network graph. Each vertex in the graph stands for a cluster. A link connects two vertices if there are gateways between these two clusters.

3. The weight of the link is the contact probability of the corresponding gateway nodes. Based on the network graph, the shortest path algorithm is employed to establish the routing table. Each entry in the routing table consists of the ID of a destination cluster and the next-hop cluster ID.

4. Once the routing table is obtained, the routing is performed from a cluster to another cluster via One-hop Inter-cluster Routing and Intra cluster Routing. The diagram for cluster based routing protocols is given below.
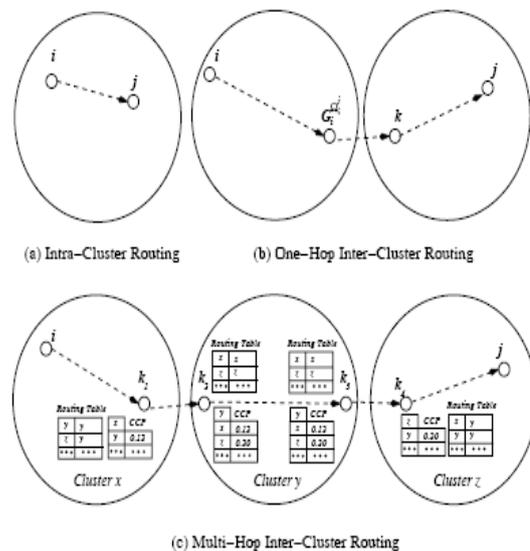
Fig 2.Cluster based routing protocol

By using the above mentioned concepts with the block diagram and the flow control the construction of Exponential Weighted Moving Average become simple. Thus the distributed clustering algorithm used to form a cluster in delay tolerant mobile networks and the cluster based routing protocol used to perform routing in delay tolerant mobile networks.

5. EXPERIMENTAL SETUP:

The performance of power balanced communication scheme is evaluated using Network simulator -2, which simulate node mobility [5], realistic physical layer radio network interface and AODV protocol. Evaluation is based on the simulation of 50 nodes located in the area of 1500 x 300 m^2. The traffic simulator is constant bit rate (CVR). The three different scheme non cluster method, EWMA, and power balanced communication are used for comparison.

PERFORMANCE METRICS

The performance metric used in this project are throughput packet delivery ratio bandwidth end to end delay, energy construction and routing overhead. The measures and details of the various parameters are given below.

A.PACKET DELIVERY RATIO

It is defined to be the percentage of the ratio of number of packets received to the number of packets sent.

$$PDR = \frac{\text{Number of packets received}}{\text{Number of packets sent}} \times 100\%$$
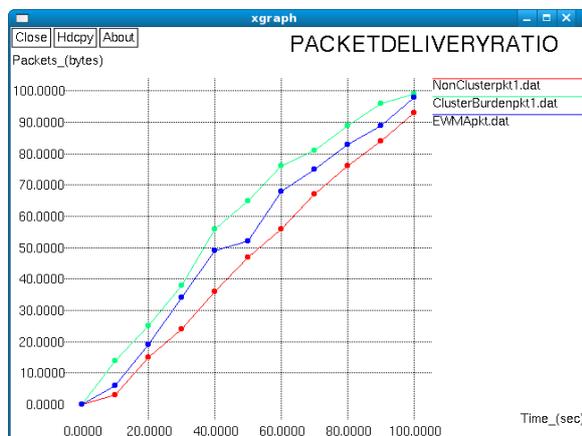
Fig2. Performance comparison using Packet delivery ratio

The X graph shows the variation of Packets (bytes) received based on the time when three different routing schemes are implemented.

C.END TO END DELAY:

The time interval between the first packet and second packet. Here the total delay takes 1.3 in non-cluster method and 0.9 in EWMA and power balanced communication have 0.4.
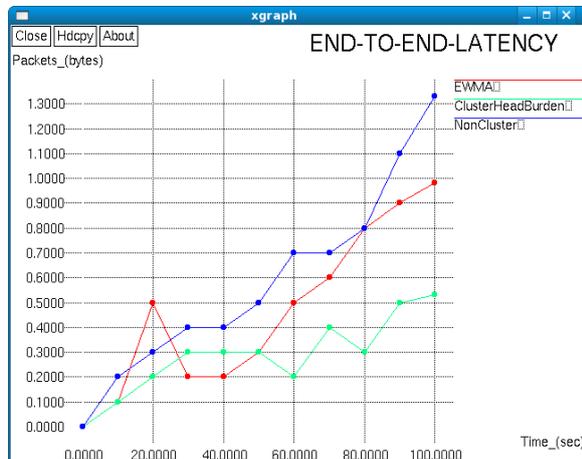
Fig3. Performance comparison using end to end latency

From the X graph shows the proposed power balance communication system achieves low end to end delay.

C. ROUTING OVERHEAD

Total number of route request and the route reply at the time
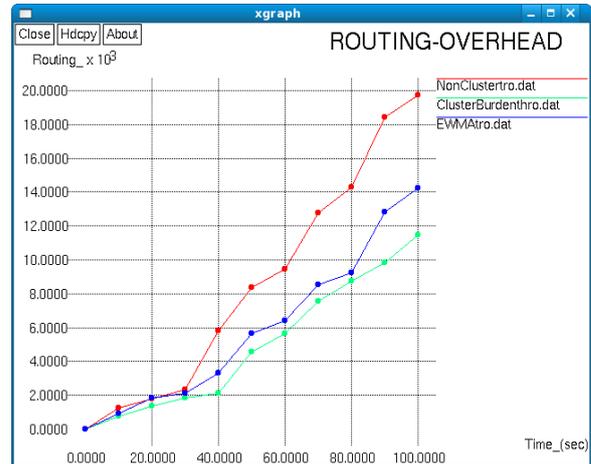
Fig4.Performance using routing overhead

From the X graph shows the routing over head has low in power balanced communication when compare to other existing methods.

D.TRHOUGHPUT:

Throughput is the ratio of number of packets received to the time seconds.

$$\text{Throughput} = \frac{\text{Number of packets received}}{\text{Time (sec)}}$$
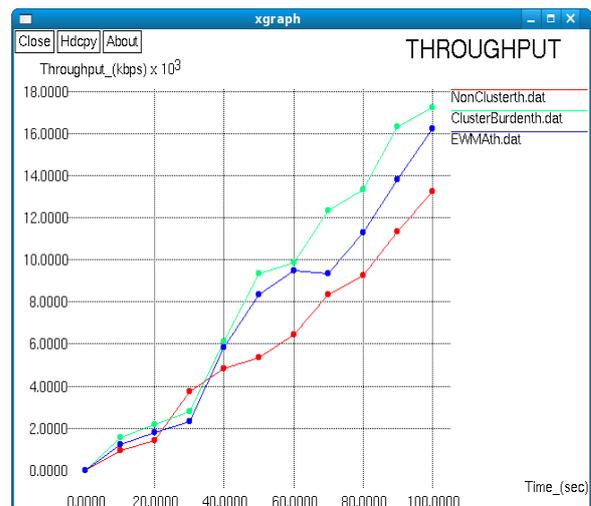
Fig5. Performance comparison using throughput

From the X graph shows the throughput high value to the Power balanced communication when compare to other existing methods.

6. CONCLUSION:

Establishing end-to-end connections for data delivery among Delay-Tolerant Mobile Networks becomes impossible as communication links only exist temporarily. In such networks, routing is largely based on nodal contact probabilities. To solve this problem, an exponentially weighted moving average (EWMA) scheme is employed for on-line updating nodal contact probability. A set of functions including sync (), leave (), and join () are devised for cluster formation and gateway selection. Finally the gateway nodes exchange network information and perform routing. The results have shown that it achieves higher delivery ratio and significantly lower overhead and end-to-end delay, compared with its non-EWMA.

7. REFERENCES:

[1]   K.Fall, 2003, A Delay –Tolerant Network architecture for challenged Internets," in Proc.ACM SIGCOMM, pp.27-34.

[2]   Y. Wang, H. Wu, F. Lin, and N.-F. Tzeng, 2008, "Cross-layer protocol design and optimization for delay/fault-tolerant mobile sensor networks," IEEE J. Sel. Areas Commun, vol. 26, no. 5, pp. 809–819.

[3]   H. Wu, Y. Wang, H. Dang, and F. Lin, 2007, "Analytic, simulation, and empirical evaluation of delay/fault-tolerant mobile sensor networks," IEEE Trans. Wireless Commun., vol. 6, no. 9, pp. 3287–3296.

[4]   Y. Wang and H. Wu, 2007, "Delay/fault-tolerant mobile sensor network (DFTMSN): a new paradigm for pervasive information gathering," IEEE Trans. Mobile Computing, vol. 6, no. 9, pp. 1021–1034.

[5]   http://www.princeton.edu/mrm/zebranet.html.

[6]   Y. Wang and H. Wu, 2006, "DFT-MSN: the delay fault tolerant mobile sensor network for pervasive information gathering," in Proc. 26th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'07), pp. 1235–1243.

[7]   M. Musolesi, S. Hailes, and C. Mascolo, 2005, "Adaptive routing for intermittently connected mobile ad hoc networks," in Proc. IEEE 6th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM), pp. 1–7.

[8]   T.Small and Z.J. Haas, 2005, "Resource and performance tradeoffs in delay tolerant wireless networks," in Proc. ACM SIGCOMM Workshop"

[9]   M. Rezaee, M. Yaghmaee, "Cluster based Routing Protocol for Mobile Ad Hoc Networks," in Proc. Receiv pp 2008 / Accept Feb 23, 2009.

[10]   K. Fall, "A delay-tolerant network architecture for challenged Internets," in Proc. ACM SIGCOMM, pp. 27–34, 2003.

ANALYSIS OF DATA MINING TECHNIQUES FOR INCREASING SEARCH SPEED IN WEB

B.Chaitanya Krishna[1],C.Niveditha[2],G.Anusha [2],U.Sindhu [2],Sk.Silar[2]
1(Assistant Professor, Dept. of Computer Science And Engineering, KL University.)
2(B.Tech Scholars, Dept. of Computer Science And Engineering, KL University.)

ABSTRACT
The World Wide Web contains an enormous amount of information, but it can be exceedingly difficult for users to locate resources that are both high in quality and relevant to their information needs. Issues that have to be dealt with are the detection of relevant information, involving the searching and indexing of the Web content, the creation of some metaknowledge out of the information which is available on the Web, as well as the addressing of the individual users' needs and interests, by personalizing the provided information and services. In this paper we discuss mainly two algorithms which increase the search engine speed.

Keywords: World Wide Web, Search Engines, Information Retrieval, PageRank, Google,HITS.

I.INTRODUCTION

Data mining nowadays plays an important role in searching the information on the web that include a high variety data types. For reaching this goal, datamining techniques for automatic discovering and extracting the web based information has been used as webmining. In this article, data mining which is a new method in retrieving the high amount of information has been introduced.

Every day, the WWW grows by roughly a million electronic pages, adding to the hundreds of millions already on-line. Because of its rapid and chaotic growth, the resulting network of information lacks of organization and structure. Moreover, the content is published in various diverse formats. Due to this fact, users are feeling sometimes disoriented, lost in that information overload that continues to expand.

II.PAGERANK ALGORITHM
2.1.Bringing Order to the Web

The citation (link) graph of the web is an important resource that has largely gone unused in existing web search engines. Maps are created containing as many as 518 million of these hyperlinks, a significant sample of the total. These maps allow rapid

calculation of a web page's "PageRank", an objective measure of its citation importance that corresponds well with people's subjective idea of importance. Because of this correspondence, PageRank is an excellent way to prioritize the results of web keyword searches. For most popular subjects, a simple text matching search that is restricted to web page titles performs admirably when PageRank prioritizes the results (demo available at google.stanford.edu). For the type of full text searches in the main Google system, PageRank also helps a great deal.
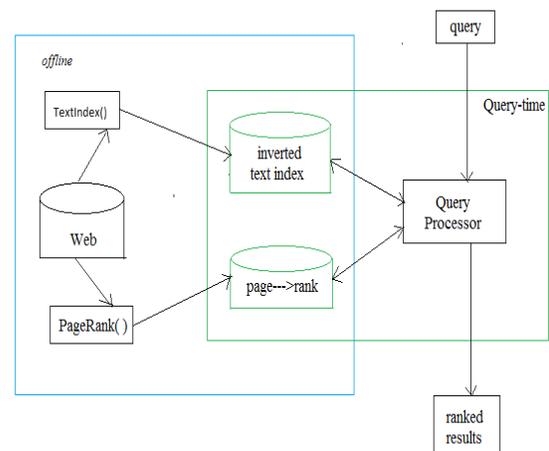
Figure 1: Simplified diagram illustrating a simple search engine utilizing the standard PageRank

Algorithm:Improved PageRank(G,s,k)

1: d <= 0.85
2: n <=number of vertices of G
3: for i = 0 to n do
4: pr[i] <= s
5: end for
6: for all E(G) do
7: t <=source(e)
8: out[i] = out[i] + 1
9: end for
10: for j = 0 to k do

11: for all E(G) = e do
12: (t,h) <= (source(e), target(e))
13: $pr_{in}[h] = pr_{in}[h] + pr[t]/out[t]$
14: end for
15: end for
16: for i = 0 to n do
17: pr[i] = (1- d) + d($pr_{in}[i]$)
18: end for
19: avg <= $\frac{sum(pr[i])}{n}$

This implementation has a complexity of
O(k m + n)

where k =number of iterations,
    G=graph
    s=integer value
    n =|V (G)|, m = |E(G)|

2.2.Description of PageRank Calculation

Academic citation literature has been applied to the web, largely by counting citations or back links to a given page. This gives some approximation of a page's importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page. PageRank is defined as follows:

We assume page A has pages T1...Tn which point to it (i.e., are citations).The parameter d is a damping factor which can be set between 0 and 1.We usually set d to 0.85. There are more details about d in the next section. Also C(A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:

PR(A) = (1-d) + d (PR(T1)/C(T1) + ... + PR(Tn)/C(Tn))

Note that the PageRanks form a probability distribution over web pages, so the sum of all web pages' PageRanks will be one. PageRank or PR(A) can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web. Also, a PageRank for 26 million web pages can be computed in a few hours on a medium size workstation.

2.3.Google Architecture Overview

Most of Google is implemented in C or C++ for efficiency and can run in either Solaris or Linux. In Google, the web crawling (downloading of web pages) is done by several distributed Crawlers. There is a URL server that sends lists of URLs to be fetched to the crawlers. The web pages that are fetched are then sent to the store server. The store server then compresses and stores the web pages into a repository. Every web page has an associated ID number called a doc ID which is assigned whenever a new URL is parsed out of a web page. The indexing function is performed by the indexer and\ the sorter. The indexer performs a number of functions. It reads the repository, uncompresses the documents, and parses them. Each document is converted into a set of word occurrences called hits. The hits record the word, position in document, an approximation of font size, and capitalization. The indexer distributes these hits into a set of "barrels", creating a partially sorted forward index. The indexer performs another important function. It parses out all the links in every web page and stores important information about them in an anchors file. This file contains enough information to determine where each link points from and to, and the text of the link. The URL resolver reads the anchors file and converts relative URLs into absolute URLs and in turn into doc IDs. It puts the anchor text into the forward index, associated with the doc ID that the anchor points to. It also generates a database of links which are pairs of doc IDs. The links database is used to compute Page Ranks for all the documents.

The URL resolver reads the anchors file and converts relative URLs into absolute URLs and in turn into docIDs. It puts the anchor text into the forward index, associated with the docID that the anchor points to. It also generates a database of links which are pairs of docIDs. The links database is used to compute PageRanks for all the documents. The sorter takes the barrels, which are sorted by docID[1] and resorts them by wordID to generate the inverted index. This is done in place so that little temporary space is needed for this operation. The sorter also produces a list of wordIDs and offsets into the inverted index. A program called DumpLexicon takes this list together with the lexicon produced by the indexer and generates a new lexicon to be used by the searcher. The searcher is run by a web server and uses the lexicon built by DumpLexicon together with the inverted index and the PageRanks to answer queries.

2.4.Google Query Evaluation

1. Parse the query.
2. Convert words into wordIDs.
3. Seek to the start of the doclist in the short barrel for every word.
4. Scan through the doclists until there is a document that matches all the search terms.
5. Compute the rank of that document for the query.
6. If we are in the short barrels and at the end of any doclist, seek to the start of the doclist in the full barrel for every word and go to step 4.

7. If we are not at the end of any doclist g o to step 4. Sort the documents that have matched by rank and return the top .

2.5.Limitations of pageRank

As we know that in PageRank the web pages are ranked according to the number of clicks made on that particular web page but this may lead to illegal ranking of web pages i.e.,whenever a query is given the pages that are satisfying the query are presented according to the rank of the page. The top most one will be given highest priority. The highest priority is because the number of clicks on that particular web page are more without concerned with the content that is present in that particular web page. For this purpose the ranking should be given according to the content present in the web page rather than the number of clicks made on that particular web page. Because if a wrong page is presented to end user then he will browse the page which will increase the click count of the traced page which is wrong. This leaves the web page with highest priority. This will continue furthurly. For this purpose it is better to rank pages according to the content in the web page. This leads to the combination of text mining with web mining.

III. HITS ALGORITHM
3.1. An overview of the HITS algorithm

HITS (Hyperlink-Induced Topic Search) algorithm mines the link structure of the Web and discovers the thematically related Web communities that consist of 'authorities' and 'hubs.' Authorities are the central Web pages in the context of particular query topics. For a wide range of topics, the strongest authorities consciously do not link to one another. Thus, they can only be connected by an intermediate layer of relatively anonymous hub pages, which link in a correlated way to a thematically related set of authorities [2].

A good hub page for a subject points to many authoritative pages on that content, and a good authority page is pointed by many good hub pages on the same subject. We should stress that a page might be a good hub and a good authority in the same time. This circular relationship leads to the definition of an iterative algorithm,HITS.

These two types of Web pages are extracted by iteration that consists of following two operations.

$$x_p = \sum_{q, q \to p} y_p$$

$$y_p = \sum_{q, p \to q} x_q$$

For a page p the weight of x_p is updated to be the sum of y_p over all pages q that link to p: where the notation $q \to p$ indicates that q links to p In a strictly dual fashion, the weight of y_p is updated to be to the sum of x_q. Therefore, authorities and hubs exhibit what could be called mutually reinforcing relationships: a good hub points to many good authorities, and a good authority is pointed to by many good hubs.
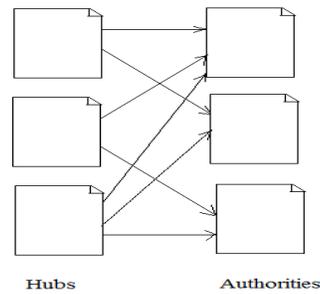
Figure 2: Illustration of Hub and Authorities

The whole picture of HITS algorithm is shown as follows:

step 1: Collect the r highest-ranked pages for the query σ from a text-based search engine such as AltaVista[3]. These r pages are referred as the root set Rσ.

step 2: Obtain the base set Sσ whose size is n by expanding $R\sigma$ to include any page pointed to by pages in $R\sigma$ and at most d pages pointing to pages in $R\sigma$.

step 3: Let G[$S\sigma$] denote the subgraph induced on the pages in $S\sigma$. Two types of links in G [$S\sigma$] are distinguished as *transverse links* and *intrinsic links*. The former are the links between pages with different domain names, and the latter are the ones between pages with the same domain name. All *intrinsic links* from the graph $S\sigma$ are deleted, keeping only the edges corresponding to *transverse links*.

step 4: Make the n by n adjacency matrix A and its transposed matrix A^T. Normalized principal eigenvector e_1 of A_T that corresponds to the largest eigenvalue λ_1 is obtained by eigenvalue calculation.

step 5: Find elements with large absolute values in the normalized principal eigenvector e_1. Return them as 'authorities.'

3.2.Pseudocode for HITS algorithm

1 G := set of pages
2 **for each** page p in G **do**
3 p.auth = 1
 // p.auth is the authority score of the page p
4 p.hub = 1 // p.hub is the hub score of the page p
5 **function** HubsAndAuthorities(G)
6 **for** step **from** 1 **to** k **do**
 // run the algorithm for k steps
7 norm = 0
8 **for each** page p in G **do**
 // update all authority values first
9 p.auth = 0
10 **for each** page q in $p.incomingNeighbors$ **do**   //
$p.incomingNeighbors$ is the set of pages that link   to
p
11   p.auth += q.hub
12   norm += square(p.auth)
// calculate the sum of the squared auth values to
normalize
13   norm = sqrt(norm)
14 **for each** page p in G **do** // update the auth scores
15   p.auth = p.auth / norm // normalize the auth
values
16   norm = 0
17   **for each** page p in G **do** // then update all hub
values
18   p.hub = 0
19   **for each** page r in $p.outgoingNeighbors$ **do**
//$p.outgoingNeighbors$ is the set of pages that p links
to
20   p.hub += r.auth
21   norm += square(p.hub) // calculate the sum of
the squared hub values to normalize
22   norm = sqrt(norm)
23   **for each** page p in G **do** // then update all hub
values
24   p.hub = p.hub / norm   // normalize the hub
values

3.3. Problems with the HITS Algorithm

To clarify problems with HITS algorithm, we traced Kleinberg's experiments. We picked 9 query topics for our study: 'abortion,' 'Artificial Intelligence,' 'censorship,' 'Harvard,' 'jaguar,' 'Kyoto University,' 'Olympic,' 'search engine,' and 'Toyota.' In these query topics, all but 'Kyoto University' and 'Toyota' were used in [2] and [3]. Though we fixed the parameters , , and a text-based search engine for collecting the root set to examine Kleinberg's experiments rigorously, we observed HITS algorithm performed poorly in several of our test cases. In this paper, we discuss focusing on topic 'Artificial Intelligence' as a successful example, and topic 'Harvard'as an unsuccessful example.

Topic: 'Artificial Intelligence'
The extracted top 5 authorities and hubs of 'Artificial Intelligence' in our experiment are indicated in Table 1. The decimal fractions shown on the left of URLs represent authority weights () and hub weights () respectively The top authority was the home page of JAIR (Journal of Artificial Intelligence Research), the second authority was AAAI (American Association for Artificial Intelligence), then MIT AI laboratory followed. Namely, famous organizations related to Artificial Intelligence based in the United States were successfully extracted. This AI community was supplemented by hubs, which consisted of the researcher's personal Web pages (e.g. S. Russell at UCB)

x_p	Authorities
.372	http://www.cs.washington.edu/research/jair/home.html
.298	http://www.aaai.org/
.294	http://www.ai.mit.edu/
.272	http://ai.iit.nrc.ca/ai_point.html
.234	http://sigart.acm.org/

y_p	Hubs
.228	http://yonezaki-www.cs.titech.ac/jp/member/hidekazu/Work/AI.html
.228	http://www.cs.berkley.edu/~russell/ai.html
.204	http://uscial.usu.clu.edu/pantonio/cco360/AIWeb.htm
.181	http://www.scms.rgu.ac.uk.staff/asga/ai/html
.171	Tpp://www.ex.ac.uk/ESE/IT/ai.html

(note: x_p and y_p represent authority weight and hub weight respectively)

Table 1: Authorities and hubs of 'Artificial Intelligence.'

Topic: 'Harvard'

In Kleinberg's experiment, authorities of 'Harvard' were related to Harvard University; e.g. the homepage of Harvard University, Harvard Law School, Harvard Business School, and so on. However, in our experiment, the Web pages authored by a financial consulting company were extracted (see Table 2). These pages did not relate to query 'Harvard.'

x_p	Authorities
.130	http://www.wetradefutures.com/investment.asp
.130	http://www.wetradefutures.com/trend.htm
.130	http://www.wetradefutures.com/market_technology.htm
.130	http://www.wetradefutures.com/florida_investment.htm
.130	http://www.wetradefutures.com/investing_investment.htm
y_p	Hubs
.247	http://www.profittaker.net/data.htm
.247	http://www.profittaker.org/new_twentyseven.htm
.247	http://www.profittaker.com/sunday_trader_more.htm
.247	http://www.profittaker.cc/system_software.htm
.247	http://www.profittaker.com/contact_phone.htm

(note: x_p and y_p represent authority weight and hub weight respectively)

Table 2: Authorities and hubs of 'Harvard.'

In this case, higher ranked 56 authorities had the same authority weights, and higher ranked 5 hubs had the same hub weights. By checking the contents of these pages, we detected that these authorities and hubs were authored by a single organization.

In the HITS algorithm, the first step is to retrieve the set of results to the search query. The computation is performed only on this result set, not across all Web pages.

Authority and hub values are defined in terms of one another in a mutual recursion. An authority value is computed as the sum of the scaled hub values that point to that page. A hub value is the sum of the scaled authority values of the pages it points to. Some implementations also consider the relevance of the linked pages.

The algorithm performs a series of iterations, each consisting of two basic steps:

- **Authority Update**: Update each node's Authority score to be equal to the sum of the Hub Scores of each node that points to it. That is, a node is given a high authority score by being linked to by pages that are recognized as Hubs for information.
- **Hub Update**: Update each node's Hub Score to be equal to the sum of the Authority Scores of each node that it points to. That is, a node is given a high hub score by linking to nodes that are considered to be authorities on the subject.

The Hub score and Authority score for a node is calculated with the following algorithm:

- Start with each node having a hub score and authority score of 1.
- Run the Authority Update Rule
- Run the Hub Update Rule
- Normalize the values by dividing each Hub score by the sum of the squares of all Hub scores, and dividing each Authority score by the sum of the squares of all Authority scores.
- Repeat from the second step as necessary.

HITS, like PageRank and Brin's PageRank, is an iterative algorithm based on the linkage of the documents on the web. However it does have some major differences:

- It is executed at query time, not at indexing time, with the associated hit on performance

that accompanies query-time processing. Thus, the *hub* and *authority* scores assigned to a page are query-specific.

- It is not commonly used by search engines.
- It computes two scores per document, hub and authority, as opposed to a single score.
- It is processed on a small subset of 'relevant' documents, not all documents as was the case with PageRank.

3.3.1 In Detail
To begin the ranking, $\forall p$, auth(p) = 1 and hub(p) = 1. We consider two types of updates: Authority Update Rule and Hub Update Rule. In order to calculate the hub/authority scores of each node, repeated iterations of the Authority Update Rule and the Hub Update Rule are applied. A k-step application of the Hub-Authority algorithm entails applying for k times first the Authority Update Rule and then the Hub Update Rule.

3.3.2.Authority Update Rule
For all p, we update auth(p) to be: $\sum_{i=1}^{n} hub(i)$ where n is the total number of pages connected to p and i is a page connected to p. That is, the Authority score of a page is the sum of all the Hub scores of pages that point to it.

3.3.3.Hub Update Rule
For all p, we update hub(p) to be: $\sum_{i=1}^{n} auth(i)$ where n is the total number of pages p connects to and i is a page which p connects to. Thus a page's Hub score is the sum of the Authority scores of all its linking pages

3.3.4.Normalization
The final hub-authority scores of nodes are determined after infinite repetitions of the algorithm. As directly and iteratively applying the Hub Update Rule and Authority Update Rule leads to diverging values, it is necessary to normalize the matrix after every iteration. Thus the values obtained from this process will eventually converge.

3.4.Further problems with the HITS algorithm:
The HITS algorithm does not always behave as expected. First, if the dominant eigenvalue of M^TM is repeated, the HITS algorithm converges to an authority vector which is *not unique*, but depends on the initial seed $\overrightarrow{a_0}$. The authority vector can be any normalized vector in the dominant eigenvalue's eigenspace. For example, for a two-level reversed

binary tree B whose edges point upwards towards the root, the eigenvalues of M^TM are 2, 2, 2, 0, 0, 0, and 0. The authority weights for the three upper nodes can be any three positive numbers that sum to 1. Second, the HITS algorithm yields *zero* authority weights for apparently important nodes of certain graphs. For example, if a leaf is added at the left middle-level node of B, then both the hub and authority weights are zero for the root and for the right half of B. We call these limitations *non-uniqueness* and *nil-weighting*, respectively.

The graphs G that are characterized are those on which the HITS algorithm is non-unique or nil-weighted. Consider an undirected graph G^1 on [n] where { i,j } is an edge of G^1 if there is a k such that (k,i) and (k,j) are directed edges of G. The HITS algorithm is non-unique or nil-weighted on G if and only if there exist i,j with positive in-degree in G such that i and j are in distinct components of G^1.

A disadvantage of the algorithm is that because it is executed at query time, it may have very poor performance, depending on the number of iterations required to rank the Hubs and authorities.

A second disadvantage of the algorithm is that it may be making assumptions about the structure of the Web that no longer hold true. In the early days, there were many Web pages that were legitimate collections of links. This was due to the poor quality of search results. With the advent of better search engine technology these collections have assumed a minor role, and have generally resolved themselves into

1- Legitimate human edited directories like www.dmoz.org

2- Link farms and directories for exploitation purposes.

3- "Link bait" pages that have useful links in order to draw links from others.

4- Pages that are themselves authoritative, but also include a number of external links, such as Wikipedia.

5- Link pages based on mutual link exchanges. Like "link bait" collections, these can be legitimate sources of authority and relevance in some cases. Political action groups are going to exchange links with like-minded groups.

The algorithm doesn't seem to take account of case number 4 - "hubs" that are also authorities, and

which are perhaps the best indicators of good external pages. On the other hand, results might be unduly influenced by link farms. It might be able to do away with the tricky identification of hubs and iterative ranking simply by relying on a number of good directories like the dmoz open directory project.

IV. COMPARATIVE ANALYIS

Algorithm	PageRank	HITS
Main Technique	Web Strucutre Mining	Web Structure Mining,Web Content Mining
Methodology	This algorithm computes the score for pages at the time of indexing of the pages.	It computes hubs and authority of the relevant pages.
Input Parameter	Black links	Content,Back,forward links
Relevancy	Less(this algorithm rank the pages on the indexing time)	More (this algorithm uses the hyper links so according to Henzinger,2001 it will give good results and also consider the content of the page)
Quality of results	Medium	Less than PR
Importance	High. Back links are considered.	Moderate. Hubs and authorities scores are utilized.
Limitation	Results come at the time of indexing and not at the query time.	Topic drift and efficiency problem.

Table 3: Summary of web page ranking algorithms

V. TOP 10 SEARCH ENGINE OPTIMIZATION TECHNIQUES

There are a lot of techniques that can be used in search engine optimization that can help or hurt you when it comes to SEO. Some of the big ones that can hurt is keyword stuffing and page cloaking. You want to always make sure that you are building a site for your visitors first and foremost. The robots aren't quite as interested in your content as you think, sorry.

You can easily make your site search engine optimized by follow 10 easy steps that will make sure everything is ready for Google to go through and understand what your pages and site are about. This can help you to get more traffic from search engines by ranking well. Once you rank well, of course, you need to make sure that you are persuading the visitors to actually click on your search listing.

5.1.Content

This is the number one for any search marketing strategy, it is impossibly important to ensure that you have content worth viewing. Without this one simply step to ensure that there is a reason for someone to be on your site, everything else is useless. There are a lot of great sites to find inspiration for writing great content that works.

5.2.Incoming Links

A link is a link is a link, but without the simplest form you aren't going to do well in search engines. The more links you have the more often you are going to be crawled. It is also important to make sure that you have the proper anchor text for your incoming links. The easiest way to gain quality links from other sites is to link to sites to let them know your site is there and hope for a reciprocal link. It is also important to make sure that you have content that is worth linking to on your site.

5.3. Web site title

Making sure that you have the right web site titles for your pages is extremely important. The keywords you place in your title are important in order to ensure that your topic is understood by Google. One of the primary factors for ranking is if the title is on-topic with the search results. Not only is it important for robots to index and understand the topic of the page either. It is important for click-through rates in the search results. Pay attention to what you click on when you are searching in Google, I know that I don't always click the first results. Using great titles and topics on your site will bring you more traffic than a number one listing. Most of the time it is within the first page, but I skim through the titles to see which looks to be more on-topic for my search query.

5.4.Heading tags

When you are laying out your site's content you have to be sure that you are creating the content flow in such a way that the heading tags are based on prominence. The most prominent of course being the h1 tag, which says "this is what this block of copy is about." Making sure you understand heading

tag structure is very important. You only want to have one (or two) h1 tags per a page. It is important to not just throw anything into an h1 tag and hope you rank for it.

5.5.Internal Linking

Making sure that your internal linking helps robots (and visitors!) to find the content on your site is huge. Using relevant copy throughout your site will tell the robots (and visitors!) more effectively what to expect on the corresponding page. You do want to make sure that on pages you don't want to rank in Google that you add a nofollow tag to ensure that the ranking flow of your site corresponds with your site's topic and interests. No one is going to be searching Google to find out what your terms of service or privacy policy are.

5.6. Keyword Density

Ensuring that you have the right keyword density for your page and sites topic is paramount. You don't want to go overboard and use the keyword every 5th word but making sure it â€œcomes upâ€• often is going to help you rank better in search engines. The unspoken rule is no more then 5% of the total copy per a page. Anymore then this and it can start to look a little spammy. Granted, you aren't shooting for 5% every time. It is really all about context and relevanceâ€" just make sure it is good, quality copy.

5.7. Sitemaps

It is always a good idea to give search engines a helping hand to find the content that is on your site. Making sure that you create and maintain a sitemap for all of the pages on your site will help the search robots to find all of the pages in your site and index them. Google, Yahoo, MSN and Ask all support sitemaps and most of them offer a great way to ensure that it is finding your sitemap. Most of the time you can simply name it sitemap.xml and the search robot will find the file effectively.

5.8.Meta Tags

Everyone will tell you that meta tags don't matter, they do. The biggest thing they matter for is click-through though. There will be a lot of times when Google will use your meta description as the copy that gets pulled with your search listing. This can help to attract the visitor to visit your web site if it is related to their search query. Definitely a much overlooked (as of late) ranking factor. Getting indexed by search engines and ranking well is just the first step. The next, and biggest, step is getting that visitor that searched for your keywords to want to click on your search listing.

5.9.URL Structure

Ensuring that your URL structure compliments the content that is on the corresponding page is pretty important. There are various methods to make this work, such as modrewrite on apache.

5.10.Domain

It can help to have keywords you are interested in ranking for within your domain, but only as much as the title, heading and content matters. One very important factor that is coming to light is that domain age is important. The older the site or domain, the better it is not spam and can do well in search results. The domain age definitely isn't a make or break factor but it does help quite a bit.

V.CONCLUSION

The www has grown into a hypertext environment of enormous complexity; and the process underlying its growth has been driven in a chaotic fashion by the individual actions of numerous participants. Our experience with hits and pagerank suggests, however, that in many respects the end product is not as chaotic as one might suppose: the aggregate behavior of user populations on the www can be studied through a mathematically clean technique for analyzing the Web's link topology, and one can use this technique to identify themes about hyperlinked communities that appear to span a wide range of interests and disciplines.

REFERENCES

[1] S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine, Computer Networks, 30(1 7): 107-117, 1998, Proceedings of the 7th International World Wide Web Conference(WWW7).

[2] J. Kleinberg. Authoritative sources in a hyperlinked environment,1997. Research Report RJ 10076 (91892), IBM.

[3] D. Gibson, J. Kleinberg, and P. Raghavan. Inferring Web communities from link topology. In Proc. 9th ACM Conference on Hypertext and Hypermedia (HyperText 98), pages 225–234, Pittsburgh PA, June 1998.

[4] http://www.altavista.com

[5] http://dustinbrewer.com/top-10-search-engine-optimization-techniques/

[6] SaekoNomura,SatoshiOyama,Tetsuo Hayamizu, Toru Ishida. Analysis and Improvement of HITS Algorithm for Detecting Web Communities. Department of Social Informatics, Kyoto University Honmachi Yoshida Sakyo-ku, Kyoto, 606-8501 Japan.

[7]   AmirPanah, AminPanah. Evaluating the datamining techniques and their roles in
Increasing the search speed data in web. 978-1-4244-5540-9/10/$26.00 ©2010 IEEE.

[8] G. Pinski, F. Narin, Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics, in Information Processing and Management. 12, (1976).

[9] D7. Gibson, J. Kleinberg, P. Raghavan, Inferring Web Communities from Link Topology, in the Proceedings of the 9th ACM Conference
on Hypertext and Hypermedia, (1998).

[10] M. Rajman, M. Vesely, From Text to Knowledge: Document Processing and Visualization: a Text Mining Approach, in Proceedings of the NEMIS Launch Conference, International Workshop on Text Mining & its Applications, Patras, Greece, April(2003).

[11] G. Salton and M. McGill. Introduction to Modern Information Retrieval. McGraw-Hill, 1983.

A Novel Approach for Link/Path Protection in Dual-Link Failures

1 Ashok kumar velpuri,2 Sreenivas velagapudi

1 Student, M.tech,Department of computer science and engineering,
K L University, Greenfields,Vaddeswaram,Guntur District,A.P,INDIA,Pincode:52250
2 Asst.prof, Department of computer science and engineering, K L University.

Abstract:

In every network we see the link failures are common, for this purpose networks having the scheme to protect their links against the link failures. Link protection helps fast recovery from link failures .Existing schemes either pre-reserve two backup paths for each demand or compute new backup paths for unprotected demands after the first link failure occurs. Both approaches require a large amount of backup capacity. In this paper, we propose a capacity efficient hybrid protection/restoration scheme for handling two-link failures. The protection component reserves backup capacity intelligently to ensure the majority of the affected demands can be restored using the pre-planned backup paths upon a two-link failure. A remarkable feature of our approach is that it is possible to trade off capacity for restorability by choosing a subset of double-link failures and designing backup paths using our algorithm for only those failure scenarios.

In this we use Backup link mutual exclusion(BLME), when the links fail simultaneously. The solution methodologies for BLME problem is 1).for mulating the backup path selection as an integer linear program;2)developing a polynomial time heuristic based on minimum cost path routing

Key Words: Optical networks, link protection, link failures, backup link mutual exclusion

I. INTRODUCTION

The growing transmission speed in the communication networks calls for efficient fault-tolerant network design. Current day's backbone networks use optical communication technology involving wavelength division multiplexing (WDM). One of the most gifted concepts for high capacity communication systems is wavelength division multiplexing (WDM). Each communication channel is allocated to a different frequency and multiplexed onto a single fiber. At
the destination wavelengths are spatially separated to different receiver locations. In this configuration the high carrier bandwidth is utilized to a greater level to transmit multiple optical signals through a single optical fiber.
            Optical networks at present operate in a circuit switched way as optical header processing and buffering technologies are still in the in the early hours stages of research for wide-scale commercial deployment. Protecting the circuits or connections established in such networks against single-link failures may be achieved in different ways:
Path protection: Path protection is having the capability to

protect one or more peer-to-peer paths via a predetermined or pre-established backup tunnel. This is for all time peer-to-peer protection and is similar to the shadow PVC model often used in the ATM networks. The backup tunnel is link and node diverged from the primary tunnel, such that if any element (link or node) along the primary path fails, the head end reroutes the traffic onto the backup path. Many schemes for backup can be used, such as 1 to N or 1 to 1. In the 1-to-N scheme, there is one backup tunnel for N primary tunnels between the same pair of routers. The 1-to-1 back up implies that for every primary tunnel a backup tunnel exists. The number of backup tunnels needed for path protection is twice the number of primary tunnels. The past is referred to as failure independent path protection (FIPP) while the latter is referred to as failure-dependent path protection (FDPP).

Link protection: As clear by the name itself, link protection involves protecting against link failures. These days, links have become more reliable, but statistics still show that most unplanned failures in the network occur because of linkfailures. So, protecting against link failures is necessary in any network. To protect against link failures it can use multiple circuits or SONET APS protected circuits. This can result in expensive circuits. Because providing circuits is usually a recurring cost especially if the fiber circuit is not owned by the carrier you might want to reduce the operating cost by eliminating the redundant circuits if fast reroute of traffic can be done by using other paths in the network. Link protection enables you to send traffic to the next hop on a backup tunnel should the primary link fail. Off-course link protection does not work if the only means of reaching the next hop is through the primary link (singly connected cases). Link protection reduces the communication requirement as compared to path protection, so providing fast recovery. On the other hand, the downside of link protection is that its capacity requirement is higher than that of path protection, explicitly when protection is employed at the connection granularity [2].

Node protection: In link protection, the backup tunnel is always set up to the next hop node and the failure detection is

performed based on loss of carrier or SONET alarms. In node protection, the mechanism described is similar to the link protection except that the backup tunnel is always set up to the node beyond the next hop that is, next-next hop. Upon detection of failure via a hello timeout, the point of local repair (PLR) node reroutes traffic onto the backup tunnel to the next-next-hop (nnhop). However, when MPLS packets emerge at the tail of the nnhop backup tunnel, they might not have the right labels for the merge point to carry the traffic further. To avoid discarding traffic at the tail of the backup tunnel, the head of the backup tunnel (also known as the point of local repair) swaps the primary tunnel label to the label expected by the merge point and then imposes the backup tunnel label. This ensures that the MPLS packets coming out of the backup tunnel carry the correct labels and hence are switched to the correct destination.

Algorithms for protection against link failures have traditionally considered single-link failures [3]–[5]. However, dual-link failures are becoming more and more important due to two reasons. First, links in the networks share resources such as conduits or ducts and the failure of such shared resources result in the failure of multiple links. Second, the average repair time for a failed link is in the order of a few hours to few days [6], and this repair time is satisfactorily long for a second failure to occur. Although algorithms developed for single-link failure resiliency is shown to cover a good percentage of dual-link failures [7]–[10], these cases often include links that are far away from each other. Considering the fact that these algorithms are not developed for dual-link failures, they may provide as an alternative to recover from independent dual-link failures. However, reliance on such approaches may not be preferable when the links close to one another in the network share resources, leading to correlated link failures.

Dual-link failures may be modeled as shared risk link group (SRLG) failures. A connection established in the network may be given a backup path under every possible SRLG failure. This approach assumes a precise knowledge of failure locations to re-configure the failed connections on their backup paths. An alternative is to protect the connections using link protection, where only the nodes adjacent to the failed link (and those involved in the backup path of the link) will perform the recovery. The focus of this paper is to protect end-to-end connections from dual-link failures using link protection.

II.DUAL-LINK FAILURE RESILIENCY WITH LINK PROTECTION

Assume that two links, l and l', failed one after the other (even if they happen together, assume that one failed first followed by the other) in a network. The backup path of the first failed link is analogous to a connection (at the granularity of a fiber) established between two nonadjacent nodes in the network with link removed. The connection is required to be protected against a single-link failure. Therefore, strategies developed for protecting connections against single link failures may be directly applied for dual-link ailures that employ link protection to recover from the first failure. Dual-link failure resiliency strategies are classified based on the nature in whichthe connections are recovered from first and second failures. The recovery from the first link failure is assumed to employ link protection strategy. Fig. 1 shows an example network where link 1-2 is protected by the backup path 1-3-4-2. The second protection strategy will refer to the manner in which the backup path of the first failed link is recovered.
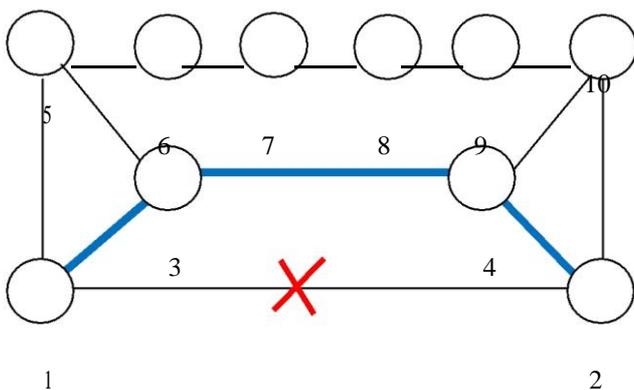
Fig. 1: Link 1-2 Protected by Backup Path 1-3-4-2 when Failed

Link Protection—Failure Independent Protection (LPFIP):
One approach to dual-link failure resiliency using link protection is to compute two link-disjoint backup paths for every link. Given a three-edge-connected network, there exists three link-disjoint paths between any two nodes [11]. Thus, for any two adjacent nodes, there exists two link-disjoint backup paths for the link connecting the two and $B'l$ denote the two link-disjoint backups for link Bl. If any link in the backup path Bl fails, the backup path of will be reconfigured to $B'l$. Hence, the nodes connected to link l must have the knowledge of the failure in its backup paths (not necessarily the location).

Link Protection—Failure Dependent Protection (LPFDP):
For every second failure that affects the backup path, a backup path under dual-link failure is provided. This backup path is

computed by eliminating the two failed links from the network and computing shortest path between the specific node pairs. When a second link failure occurs, a failure notification must be sent to node specific node. It is fairly straight forward to see that the average backup path length under dual-link failures using LP-FDP will be lesser than that using LP-FIP. Every link is assigned one backup path for single link failure and multiple backup paths (depending on the number of links in the backup path for the single link failure) under dual-link failures.

Link Protection—Link Protection (LP-LP*)*: Notification of the second failed link to different nodes for them to reconfigure their backup paths may result in a high recovery time. In order to avoid notification to the other nodes and reconfiguring at the end of the paths, link protection may be adopted to recover from the second link failure as well. Under this strategy, every link will have only one backup path (for all failure scenarios). In order for this strategy to work, the backup path under the second failure must not pass through the first failed link. This condition is referred to as the *backup link mutual exclusion* (BLME) constraint.

III. HEURISTIC APPROACH

As ILP solution times for large networks may be prohibitively high, a heuristic approach is also developed. The heuristic solution is based on iterative computation of minimum cost routing. The network is treated as an undirected graph G. A set of auxiliary graphs corresponding to failure of a link l G is created. In each auxiliary graph Zl the objective is to obtain a path between the nodes that were originally connected by link l. Let Pl denote the path selected in auxiliary graph Zl. If a link l' is a part of the path selected on graph Zl, then the path in graph Zl must avoid the use of link l. This is accomplished by imposing a cost on the links in the auxiliary graphs and having the path selection approach select the minimum cost path. Let Wll' denote the cost of link l' on graph Zl such that it indicates that graph Zl' contains link l and the two links l and l' may be unavailable simultaneously. Hence, the cost values are binary in nature.

The cost of a path in an auxiliary graph is the sum of the cost of links in it. At any given instant during the computation, the total cost of all the paths (T) is the sum of the cost of the paths across all auxiliary graphs. It may be observed that the total cost must be an even number, as every link l' in a path Pl that has a cost of 1 implies that link l in path Pl' would also have a cost of 1. For a given network, the minimum value of the total cost would then be two times the number of dual-link failure scenarios that would have the network scenarios that would disconnect the graph, then the termination condition for the heuristic is given by T= 2 T. disconnected. If T denotes the number of dual-link failure

Steps involved in the IMCP heuristic solution.

Iterative Minimum Cost Path (IMCP) Heuristic:

Step1. Obtain auxiliary graphs Zl for every l € Z as Zl =Z - {l} Note that every link l   Z is bidirectional in nature.

Step 2. Initialize the path to be found in every graph Zl as an empty set Pl ← 0 , l€ Z

Step 3. Initialize the cost of all the links in every auxiliary graph to 0, Wll'← 0, l€ Z, l'   Zl

Step 4. For every auxiliary graph Zl

 1. Erase the old path and and update the cost in auxiliary graphs; ie, for every link l' € Pl update Wl'l ← 0,   Pl ← 0

 2. Recompute the least cost path Pl

 3. If the link l' is present in this graph, then modify the cost of link l in auxiliary graph Zl, ie for every link l'   Pl
 update   Wl'l ← Pll'

 Step 5.Compute the total cost of all path over all the auxiliary graphs ie T = Σ l Z Σ l'   Pl Wll'

Step 6.If the total cost all the paths equals the threshold of 2T , where T is the number of dual link failure scenarios that would disconnect the graph, then it indicated the best possible solution has been obtained, ie T= 2 T, go to step 7, otherwise go to step 4.
Step 7:stop

IV. SYSTEM ANALYSIS
Existing System:
 Algorithms for protection against link failures have traditionally considered Single-link failures. However, dual link failures are becoming increasingly important due to two reasons. First, links in the networks share resources such as conduits or ducts and the failure of such shared resources result in the failure of multiple links. Second, the average repair time for a

failed link is in the order of a few hours to few days, and this repair time is sufficiently long for a second failure to occur. Algorithms developed for single-link failure resiliency is shown to cover a good percentage of dual-link failures, these cases often include links that are far away from each other. Considering the fact that these algorithms are not developed for dual-link failures, they may serve as an alternative to recover from independent dual-link failures.

Proposed System:

This paper formally classifies the approaches for providing dual-link failure resiliency. Recovery from a dual-link failure using an extension of link protection for single link failure results in a constraint, referred to as BLME constraint, whose satisfiability allows the network to recover from dual-link failures without the need for broadcasting the failure location to all nodes.
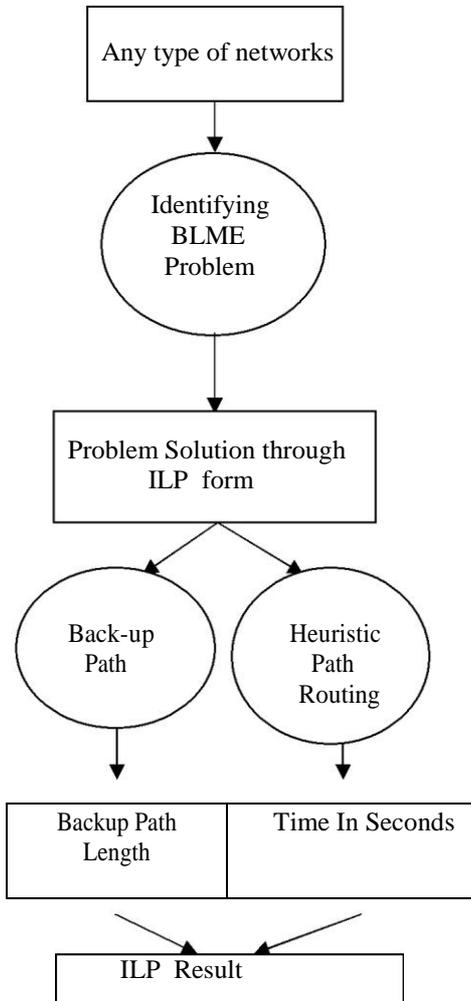
Fig. 2: Flow Diagram

This paper develops the necessary theory for deriving the sufficiency condition for a solution to exist, formulates the problem of finding backup paths for links satisfying the BLME constraint as an ILP, and further develops a polynomial time heuristic algorithm. The formulation and heuristic are applied to six different networks and the results are compared.

V. MODULE DESCRIPTION

Module 1:
containment hierarchy is a tree of components that has a top-
Component creation: To appear onscreen, every GUI component must be part of a *containment hierarchy.*

A level container as its root. Each GUI component can be contained only once. If a component is already in a container and try to add it to another container, the component will be removed from the first container and then added to the second.

Module 2:
Application of events and positioning of components

• Create the nodes in different positions and apply different colors.
• Create the distance between the nodes by applying stress.
• Apply different mouse events to the nodes.
• Using grouplayout of JFreechart all the nodes are positioned. By using virtual and horizontal position and parallel group, all the nodes are positioned.

Module 3:
Calculating TIS and BP length

• By clicking any node, the data transfer to the next node and dual failure are shown in red color.
• For this dual failure, the backup path is shown in red color.

• The Time in seconds and BP length is displayed.

In this paper heuristic are applied only for six different types of networks [11] that are shown here.

(a)   (b)

(b)   (d)
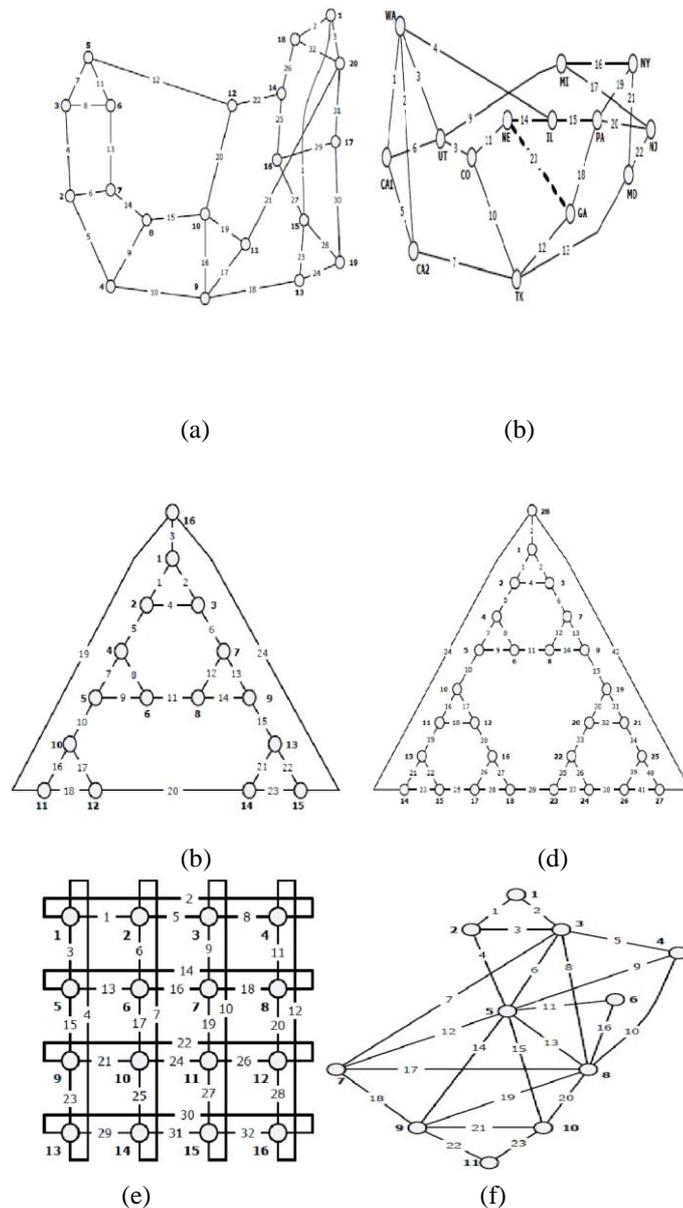
(e)   (f)
Fig. 3: Networks Considered for Performance Evaluation.

(a) ARPANET (20 nodes, 32 links). (b) NSFNET (14 nodes, 23 links). (c) Node-16 (16 nodes, 24 links). (d) Node-28 (28 nodes, 42 link nodes, 23 links).

CONCLUSION

This paper focuses on the approaches for providing dual-link failure resiliency. Recovery from a dual-link failure using an extensio constraint, referred to as BLME constraint, whose satisfiability allows the network to recover from dual-link failures without the need paper develops the necessary theory for deriving the sufficiency condition for a solution to exist, heuristic is shown to obtain a so guarantee, although such a solution may have longer average hop lengths compared with the optimal values. The heuristic produces a so the six scenarios. A maximum of 30 iterations were performed. While the objective of the heuristic is to obtain a feasible solution, it is n network scenario for any arbitrary two link failure scenario). The number of iterations required to arrive at the solution depends on a auxiliary graphs are considered and the weights employed. Comparing the results of the heuristic to that of the ILP, it is observed tha average backup path lengths.

References:
[1]   A. Chandak and S. Ramasubramanian, "Dual-link Failure Resiliency through Backup Link Mutual Exclusion," *in Proc. IEEE Int. Conf. Broadband Networks*, Boston, MA, Oct. 2008, pp. 258–267.
[2]   J. Doucette and W. D. Grover, "Comparison of Mesh Protection and Restoration Schemes and the Dependency on Graph Connectivity," Hungary, Oct. 2001, pp. 121–128.
[3]   M. Medard, S. G. Finn, and R. A. Barry, "WDM Loopback Recovery in Mesh Networks," *in Proc. IEEE INFOCOM*, 2008, pp. 752–759.
[4]   S. S. Lumetta, M. Medard, and Y. C. Tseng, "Capacity Versus Robustness: A Tradeoff for Link Restoration in Mesh Networks," *J. Lightw. Technol.*, 18, No. 12, pp. 1765–1775, Dec. 2000.
[5]   G. Ellinas, G. Halemariam, and T. Stern, "Protection Cycles in WDM Networks," *IEEE J. Sel. Areas Commun.*, 8, No. 10, pp. 1924–1937, Oct. 2000.
[6]   W. D. Grover, *Mesh-Based Survivable Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking*. Upper Saddle River, NJ: Prentice-Hall, 2007. M. Fredrick, P. Datta, "Sub-graph Routing: A Novel Fault-tolerant Architecture for Sharedrisk Link Group Failures
[7]   M. Clouqueur and W. D. Grover, "Mesh-restorable Networks with Complete Dual-failure Restorability and with Selectively Enhanced Dual-failure Restorability Properties," in *Proc. OPTICOMM*, 2002, pp. 1–12.
[8]   J. Doucette and W. D. Grover, "Shared-risk Logical San Groups in Span-restorable Optical Networks:Analysis and Capacity Planning Model," *Photon. Netw. Commun.*, 9, No. 1, pp. 35–53, Jan. 2005.
[9]   J. A. Bondy and U. S. R. Murthy, *Graph Theory With Applications*. New York: Elsevier, 2008.
[11] H. Choi, S. Subramaniam, and H. Choi, "On Double-link Failure Recovery in WDM Optical Networks," in *Proc. IEEE INFOCOM*, 2007, pp. 808–816.
[12] CPLEXSolver.[Online].Available:http://www.cplex.com
[13] H. Choi, S. Subramaniam, and H. Choi, "Loopback Recovery from Double-link Failures in Optical Mesh Networks," *IEEE/ACM Trans. Netw.*, 12, No. 6, pp. 1119– 1130, Dec. 2004.
[14]   H. Choi, S. Subramaniam, and H.-A. Choi, "Loopback Recovery from Neighboring Double-link Failures in WDM Mesh Networks," *Inf. Sci. J.*, 149, No. 1, pp. 197– 209, Jan. 2003

Magnetohydrodynamic Rayleigh Problem with Hall Effect

Haytham Sulieman[(1)], Naji A. Qatanani[(2)]

(1) Palestine Polytechnic University, Hebron, Palestine
(2) Department of Mathematics, An-Najah National University,Nablus, Palestine

Abstract

This paper gives very significant and up-to-date analytical and numerical results to the magnetohydrodynamic flow version of the classical Rayleigh problem including Hall effect. An exact solution of the MHD flow of incompressible, electrically conducting, viscous fluid past a uniformly accelerated and insulated infinite plate has been presented. Numerical values for the effects of the Hall parameter **N** and the Hartmann number **M** on the velocity components **u** and **v** are tabulated and their profiles are shown graphically. The numerical results show that the velocity component u increases with the increases of **N** and decreases with the increase of **M**, whereas, the velocity component v increases with the increase of both **M** and **N**. These numerical results have shown to be in a good agreement with the analytical solution.

Keywords: MHD flow, Hall effect, viscous fluid, accelerated plate.

AMS 2000 mathematics subject classification numbers: 76D05, 76W05, 80A20.

1. Introduction

When the magnetic field diffuses easily through the conducting medium and when the frequency of collision of charge particles is large compared to their frequency of rotation about the magnetic field lines, the current in the medium is controlled by the resistance of the medium and in such a case the generalized Ohm's law is the appropriate law to apply . However, if these conditions are not fulfilled, additional terms will appear in the generalized Ohm's law.

The MHD Stokes or Rayleigh problem was first solved by Rossow [11] without taking into account the induced magnetic field. With the induced magnetic field, it was solved by Nanda & Sundaram [8], Ludford [6], Chang & Yen [2], Rosciszewski [10], and Hashimoto [4]. In these papers, different aspects of the problem were considered. But in an ionized gas where the density is low and / or the magnetic field is very strong, the conductivity becomes a tensor. The conductivity normal to the magnetic field is reduced by the free spiraling of electrons and ions about the magnetic lines of force before they experience collisions, and a current, known as the Hall current is induced in a direction normal to both electric and magnetic fields. Steady state channels flows of ionized gases were studied by Sato [12], Yamanishi [17] and Sherman & Sutton [14]. The effects of Hall current on MHD Rayleigh's problem in ionized gas where studied by Mohanty [7]. Schlichting [13] has studied the unsteady flow due to an impulsive motion of an infinite plate in a fluid of an infinite extent. He showed that this simple problem admitting an exact solution for the Navier-Stokes equation. MHD flow past a uniformly accelerated plate under a transverse magnetic field was studied by Gupta [3], Pop [9] and Soundalgekar [15], neglecting induced magnetic field. Kinyanjui [5] studied the heat and the mass transfer in unsteady free convection flow with radiation absorption passed an impulsively started infinite vertical porous plate subjected to strong magnetic field including the Hall effect. Maleque and Sattar [1] investigated the steady laminar flow on a porous rotating disk with variable fluid properties taking Hall effect into account.

The study of the MHD flow with Hall currents has important engineering applications in problems of MHD generators. Hall accelerators as well as in flight Magnetohydrodynamics. The rotating flow of an electrically conducting fluid in the presence of magnetic field is encountered in cosmical and geophysical fluid dynamics. It is also important in the solar physics involved in the sunspot development, the solar cycle and the structure of rotating magnetic stars.

In this study we have considered the effect of the Hall current on the magnetohydrodynamics flow version of the classical Rayleigh problem. Thus, an exact solution of the MHD flow of incompressible, electrically conducting and insulated infinite plate has been presented. Numerical results for the effects of the Hall parameter **N** and the Hartmann number **M** on the velocity components **u** and **v** are tabulated and their profiles are shown graphically.

2. Formulation of the problem

We consider the flow of an incompressible, electrically conducting, viscous fluid past an infinite and insulated flat plate occupying the plane $y = 0$. Let the positive direction of x-axis be chosen along the plate in the direction of the flow and the y-axis is normal to it. A uniform magnetic field H_0 is applied in the direction of the y-axis. The physical configuration and the nature of the flow suggest the following form of velocity vector $\vec{q}$, magnetic induction vector $\vec{H}$, electrostatic field $\vec{E}$ and pressure P, thus:

$$\left.\begin{array}{l} \vec{q} = (\, u, 0, v \,) \\[4pt] \overline{H} = (\, H_x, H_0, H_z \,) \\[4pt] \vec{E} = (\, E_x, 0, E_z \,) \\[4pt] P = \text{constant} \end{array}\right\} \tag{2.1}$$

The equations governing the unsteady flow and Maxwell's equations are:

Equation of continuity:

$$\nabla . q = 0 \tag{2.2}$$

Equation of motion:

$$\frac{\partial q}{\partial t} + (q.\nabla) \, q = - \frac{1}{\rho} \, \nabla p + \nu \, \nabla^2 \, q + \frac{1}{\rho} \, J \times H \tag{2.3}$$

Equation for current:

$$\nabla \times H = \mu \, J \tag{2.4}$$

Faraday's Law:

$$\nabla \times E = - \frac{\partial H}{\partial t} \tag{2.5}$$

$$\nabla . H = 0 \tag{2.6}$$

The generalized Ohm's law, neglecting ion-slip effect but taking Hall current is,

$$\frac{J}{\sigma} = (E + \vec{q} \times H) - \frac{(J \times H)}{n_e . e} \tag{2.7}$$

Where $\sigma = \dfrac{e^2 \, \tau \, n}{m}$ (is the electrical conductivity).

Here J is the current density, t is the time, ρ, ν, and μ stand for the density, the kinematics viscosity, and the magnetic permeability, e and m are the electric charge and the mass of an electron, n is the electron number density and τ is the mean collision time.

The Lorentz force per unit volume is given by:

$$\vec{J} \times \vec{H} = [\, - J_z H_0, J_z H_x - J_x H_z, J_x H_0] \tag{2.8}$$

Moreover:

$$\vec{q} \times \vec{H} = [\, - vH_0, vH_x - uH_z, uH_0 \,] \tag{2.9}$$

where: $\vec{J} = [\, J_x, \, 0, \, J_z \,]$

with

$$J_x = \frac{\sigma}{1 + \omega^2 \, \tau^2} \, [\, E_x - vH_0 + \omega\tau \, (E_z + uH_0) \,] \tag{2.10}$$

$$J_z = \frac{\sigma}{1 + \omega^2 \, \tau^2} \, [\, E_z - uH_0 - \omega\tau \, (E_x - vH_0) \,] \tag{2.11}$$

where $\omega = \dfrac{eH_0}{m}$ (is the electron Larmor frequency).

The initial and boundary conditions are:

$$\left.\begin{array}{l} t \leq 0: \quad u = 0, \qquad v = 0 \quad \text{for } y \geq 0 \\[4pt] t > 0: \quad u = U_0, \; v = 0 \quad \text{for } y = 0 \\[4pt] u \to 0: \; v = 0 \qquad \text{as } y \to \infty \\[4pt] H_x \to 0 \; H_y = H_0, \quad H_z \to 0 \quad \text{as } y \to \infty \end{array}\right\} \tag{2.12}$$

At infinity the magnetic induction is uniform with components $(0, H, 0)$, and hence the current density in (2.4) vanishes. And since the free stream is at rest, it follows from generalized Ohm's low that $E = 0$ as

$y \rightarrow \infty$. Assuming small magnetic Reynolds number for the flow, the induced magnetic field is neglected in comparison to the applied constant field H_0.

On introducing the non-dimensional quantities:

$$y^* = \frac{U_0 \cdot y}{\nu}, \ u^* = \frac{u}{U_0}, \ v^* = \frac{v}{U_0}, \ t^* = \frac{U_0^2 \ t}{\nu} \qquad (2.13)$$

then, we can write $u \, (y, t) = u^* \, (y^* , t^*)$

$$\frac{\partial u}{\partial t} \, (y, t) = \frac{\partial u^*}{\partial t^*} \, (y^*, t^*) = \frac{U_0^2}{\nu} \ \frac{\partial u^*}{\partial t^*} \quad \text{and} \quad \frac{\partial^2 u}{\partial y^2} = \frac{U_0^2}{\nu^2} \ \frac{\partial^2 u^*}{\partial y^{*2}}$$

in this case we obtain :

$$\frac{1}{\rho} (J \times H) = \frac{-1}{\rho} (J_z \, H_0)$$

or $\qquad \dfrac{1}{\rho} (J \times H) = - \ \dfrac{1}{1 + \omega^2 \tau^2} \ [\ \dfrac{\sigma H_0^2}{\rho} \ (u + \omega \tau v) \ ] \qquad (2.14)$

Likewise, $v \, (\, y, t \,) = v^* (\, y^*, t^* \,) \qquad$ yields

$$\frac{\partial v}{\partial t} = \frac{U_0^2}{\nu} \ \frac{\partial v^*}{\partial t^*} \quad \text{and} \quad \frac{\partial^2 v}{\partial y^2} = \frac{U_0^2}{\nu} \ \frac{\partial^2 v^*}{\partial y^{*2}}$$

Then we obtain: $\dfrac{1}{\rho} (J \times H) = \dfrac{1}{\rho} (J_x \, H_0)$

or $\qquad \dfrac{1}{\rho} (J \times H) = \dfrac{1}{1 + \omega^2 \tau^2} \ [\ \dfrac{\sigma H_0^2}{\rho} \ (\omega \tau u - v) \ ] \qquad (2.15)$

Consequently, the equation of motion (2.3) in component term becomes
(dropping the stars):

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial y^2} - \frac{\sigma H_0^2 \ \nu}{\rho U_0^2 + \rho U_0^2 \ \omega^2 \tau^2} \ (u + \ \omega \tau v) \qquad (2.16)$$

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial y^2} + \frac{\sigma H_0^2 \ \nu}{\rho U_0^2 + \rho U_0^2 \ \omega^2 \tau^2} \ (\omega \tau v - v) \qquad (2.17)$$

Now let $M^2 = \dfrac{\sigma H_0^2 \ \nu}{\rho U_0^2}$ is the Hartmann number and $N = \omega \tau$ is the

Hall parameter **u** and **v** are the velocity components in the **x** and **y** direction respectively. The initial and boundary conditions become:

$$\left. \begin{array}{l} u(\, 0, y \,) = v(\, 0, y \,) = 0 \\ u(\, t, 0 \,) = 1, v(\, t, 0 \,) = 0 \\ u(\, t, y \,) \ \text{and} \ v(\, t, y \,) \rightarrow 0 \ \text{as} \ y \rightarrow \infty \end{array} \right\} \qquad (2.18)$$

Now, multiply both sides of equations (2.16) and (2.17) by e^{-st} and integrate from 0 to ∞ with respect to t we get:

$$\frac{d^2 \hat{u}}{d y^2} - (\ \frac{M^2}{1 + N^2} + s \ ) \ \hat{u} = \frac{NM^2}{1 + N^2} \ \hat{v} \qquad (2.19)$$

$$\frac{d^2 \hat{v}}{d y^2} - (\ \frac{M^2}{1 + N^2} + s \ ) \ \hat{v} = - \ \frac{NM^2}{1 + N^2} \ \hat{u} \qquad (2.20)$$

where:

$$\hat{u}\,(\,s,y\,) = L\,\{\,u(\,t,y\,)\,\} = \int_0^\infty u(\,t,y\,)\,e^{-st}\,dt$$

$$\hat{v}\,(\,s,y\,) = L\{\,v(\,t,y\,)\,\} = \int_0^\infty v(\,t,y\,)\,e^{-st}\,dt$$

By introducing the complex function $\hat{q} = \hat{u} + i\,\hat{v}$, then equations (2.19) and (2.20) can be combined into the single equation:

$$\frac{d^2\,\hat{q}}{d\,y^2} - \{\,\frac{M^2}{1+N^2}\,(\,1-iN) + s\,\}\,\hat{q} = 0 \tag{2.21}$$

3. Analytical Solution

By introducing the complex function $\hat{q} = \hat{u} + i\,\hat{v}$, then equations (2.16) and (2.17) yield

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial y^2} - (\,\frac{M^2}{1+N^2}\,)\,(\,1-iN)\,q \tag{3.1}$$

The initial and boundary conditions take the form:

$$\left. \begin{array}{l} q(\,0,y\,) = 0, \quad q(\,t,0) = 1 \\ q(\,t,y\,) \to 0 \quad \text{as} \quad y \to \infty \end{array} \right\} \tag{3.2}$$

Using the abbreviation $\alpha = -\dfrac{M^2}{1+N^2}\,(\,1-iN\,)$, equation (3.1) can be written as:

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial y^2} + \alpha\,q \tag{3.3}$$

let:

$$\Phi(\,t,y\,) = e^{-\alpha t}\,q(\,t,y\,) \tag{3.4}$$

and multiplying equation (3.3) by ($e^{-\alpha t}$) we get:

$$\frac{\partial\,\Phi}{\partial\,t} = \frac{\partial^2\,\Phi}{\partial\,y^2} \tag{3.5}$$

From equations (3.2) and (3.4) we conclude that:

$$\left. \begin{array}{l} \Phi(\,0,y\,) = 0, \quad \Phi(\,t,0\,) = e^{-\alpha t} \\ \Phi(\,t,y\,) \to 0 \quad \text{as} \quad y \to \infty \end{array} \right\} \tag{3.6}$$

To solve (3.5) subject to the initial and the boundary conditions (3.6) we apply the Laplace transform method.

$$L\{\,\frac{\partial\,\Phi}{\partial\,t}\,\} = L\{\,\frac{\partial^2\,\Phi}{\partial\,y^2}\,\}$$

$$s\,\hat{\Phi}(\,s,y\,) = \frac{d^2\,\hat{\Phi}}{dy^2} \tag{3.7}$$

where:

$$\hat{\Phi}(\,s,y\,) = L\{\,\Phi(t,y)\,\}$$

$$\hat{\Phi}(\,s,0\,) = \frac{1}{s+\alpha}$$

$$\lim_{y \to \infty} \hat{\Phi}(s, y) = 0 \tag{3.8}$$

The auxiliary equation for equation (3.7) can be written as:

$$\beta^2 - s = 0 \tag{3.9}$$

hence,

$$\hat{\Phi}(s, y) = C_1\, e^{-\sqrt{s}\, y} + C_2\, e^{\sqrt{s}\, y} \tag{3.10}$$

We claim $C_2 = 0$

Proof of claim:

divide both sides of equation (3.10) by $e^{\sqrt{s}\, y}$

$$e^{-\sqrt{s}\, y}\, \hat{\Phi}(s, y) = C_1\, e^{-2\sqrt{s}\, y} + C_2$$

Now ,taking the limit of both sides of the above equation as $y \to \infty$:

$$0.0 = C_2 + C_1 \cdot 0 \quad \text{i.e. } C_2 = 0$$

Furthermore:

$$\hat{\Phi}(s, y) = C_1\, e^{-\sqrt{s}\, y} \tag{3.11}$$

Setting $y = 0$ in equation (3.11) and from equation (3.8), we obtain:

$$\hat{\Phi}(s, 0) = C_1 \cdot 1 = \frac{1}{s + \alpha} \quad \text{i.e. } C_1 = \frac{1}{s + \alpha}$$

This gives

$$\hat{\Phi}(s, y) = \frac{1}{s + \alpha}\, e^{-\sqrt{s}\, y} \tag{3.12}$$

Taking the inverse transform
We have

$$\Phi(t, y) = L^{-1} \{ \hat{\Phi}(s, y) \} = L^{-1} \left\{ \frac{s}{s + \alpha} \cdot \frac{e^{-\sqrt{s}\, y}}{s} \right\} \tag{3.13}$$

Now, we use the following fact (convolution theorem) about Laplace transformation:

$$L\left\{ \int_0^t f(t - \tau)\, g(\tau)\, d(\tau) \right\} = L\{ f(t) \} \cdot L\{ g(t) \} = \hat{f}(s)\, \hat{g}(s)$$

where:

$$f(t) = L^{-1}\left\{ 1 - \frac{\alpha}{s + \alpha} \right\} = \delta(t) - \alpha\, e^{-\alpha t}$$

$$g(t) = L^{-1}\left\{ \frac{e^{-\sqrt{s} y}}{s} \right\} = \operatorname{erfc}\left(\frac{y}{2\sqrt{t}} \right) = \frac{2}{\sqrt{\pi}} \int_{\frac{y}{2\sqrt{t}}}^{\infty} e^{-u^2}\, du$$

thus, $\Phi(t, y) = L^{-1}\left\{ \dfrac{s}{s+\alpha} \cdot \dfrac{e^{-\sqrt{s} y}}{s} \right\} = \displaystyle\int_0^t [\, \delta(t - \tau) - \alpha\, e^{-\alpha(t-\tau)}\,] \times \left[\frac{2}{\sqrt{\pi}} \int_{\frac{y}{2\sqrt{\tau}}}^{\infty} e^{-u^2} du \right] d\tau$

$$\Phi(t, y) = \frac{2}{\sqrt{\pi}} \int_{\frac{y}{2\sqrt{t}}}^{\infty} e^{-u^2}\, du - \frac{2\alpha\, e^{-\alpha t}}{\sqrt{\pi}} \int_0^t \left(e^{\alpha\tau} \int_{\frac{y}{2\sqrt{\tau}}}^{\infty} e^{-u^2}\, du \right) d\tau \tag{3.14}$$

Recall,

$$q(t, y) = e^{\alpha t}\, \Phi(t, y)$$

Then we get:

$$q(t, y) = e^{\alpha t} \operatorname{erfc}\left(\frac{y}{2\sqrt{t}} \right) - \alpha \int_0^t e^{\alpha\tau} \operatorname{erfc}\left(\frac{y}{2\sqrt{\tau}} \right) d\tau \tag{3.15}$$

where

$$\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-u^2} \, du$$

Now, writing $q(\,t, y\,)$ as $u + iv$

We have,

$$q(\,t, y\,) = e^{at} \cos bt \; \text{erfc}(\frac{y}{2\sqrt{t}}) - \int_{0}^{t} e^{a\tau} \, \text{erfc}(\frac{y}{2\sqrt{\tau}})$$

$$[a \cos(b\tau) - b \sin(b\tau)] \, d\tau + i[\, e^{at} \sin(bt) \, \text{erfc}(\frac{y}{2\sqrt{t}})$$

$$- \int_{0}^{t} e^{a\tau} \, \text{erfc}(\frac{y}{2\sqrt{\tau}}) \, [a \sin(b\tau) + b \cos(b\tau)]] \, d\tau \qquad (3.16)$$

where: $\alpha = a + ib$

with

$$a = -\frac{M^2}{1 + N^2}, \quad b = \frac{M^2 N}{1 + N^2}.$$

4. Numerical Solution for The Second Order BVP

In order to get a clear understanding of the flow fluid we have carried out numerical calculations of equation (2.21). The boundary-value problem can be stated as:

$$\frac{d^2 \, \hat{q}}{d \, y^2} - \omega \, \hat{q} = 0 \qquad (4.1)$$

$$\hat{q}(\,0, s\,) = \frac{1}{s}, \quad \hat{q}(\,\infty, s) = 0 \qquad (4.2)$$

where:

$$\omega = (\frac{M^2}{1 + N^2} + s) - i \, \frac{N \, M^2}{1 + N^2}.$$

To ensure that the Laplace Transforms are well-defined, it should be assumed that $s > 0$.   This implies $\text{Re}(\omega) = \frac{M^2}{1 + N^2} + s > 0$. Hence there exists η in the complex number such that $\eta^2 = \omega$ with $\text{Re}(\eta) < 0$

Furthermore,

$$\hat{q}(\,y, s\,) = \frac{e^{\eta y}}{s} \qquad (4.3)$$

Satisfies the boundary value problem (4.1) -(4.2)   .

For **y = 0** we have:

$$\hat{q}(\,0, s\,) = \frac{1}{s} = \int_{0}^{\infty} 1 \, . \, e^{-st} \, dt = \int_{0}^{\infty} (1 + 0i) \, e^{-st} \, dt.$$

Thus,

$$u(\,0, t\,) \equiv 1 \; \text{ and } \; v(\,0, t\,) \equiv 0 \; \text{ for all } \; t$$

Recall that the inverse Laplace Transform is:

$$q(\,y, t\,) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \hat{q}(\,y, s\,) \, e^{st} \, ds$$

Where $\gamma > 0$ is chosen so that all the singularities of $\hat{q}(\,y, s\,)$ are to the left of γ. The above integral is over the vertical line $z = \gamma$ in the complex plane. Since $\hat{q}(\,y, s\,) = \frac{e^{\eta y}}{s}$, we can choose γ to be any positive

number. In the calculations below we choose $\gamma = 0.25$. We will define q strictly as a function of t using Mathematics' NIntegrate command. We will approximate the integral above by integrating from $0.25 - 500\,i$ to $0.25 + 500\,i$.

We also define $\omega = (\dfrac{M^2}{1 + N^2} + s) - i\,\dfrac{N\,M^2}{1 + N^2}$,

where $M^2 = \dfrac{\sigma\,H_0^{\,2}\,\nu}{\rho\,U_0^{\,2}}$ is the Hartmann number, and $N = \omega\tau$ is the Hall parameter.

Next, we will define the range of t values as required in cases (**1.a-4.a**)
$t = \{0.4\,,0.8\,,1.2\,,1.6\,,2\,,3\,,4\,,5\,,6\,,7\,,8\,,9\,\}$.

5. Numerical Results

The effect of the Hall parameter **N** and the Hartmann number **M** in the velocity components **u** and **v** is illustrated in the following cases:

Case 1.a

In this case, M= 1, N = 1/2, and y = 0. This case has already been eliminated since y=0. Hence this is the perfect opportunity to check the numerical method that will be employed in the other cases.

Next, we obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0.4	0.9992	0
0.8	1.0005	0
1.2	1.0007	0
1.6	1.0003	0
2	0.9997	0
3	1	0
4	1.0002	0
5	0.9997	0
6	1.0005	0
7	0.9995	0
8	1.0004	0
9	0.9998	0

Table (1): The velocity components u and v for different values of t

As the above table indicates $u(t) = \mathrm{Re}(q(t)) \approx 1$ and $v(t) = \mathrm{Im}(q(t)) \approx 0$ for all t. this is consistent with the exact results that we obtained for the case y = 0.
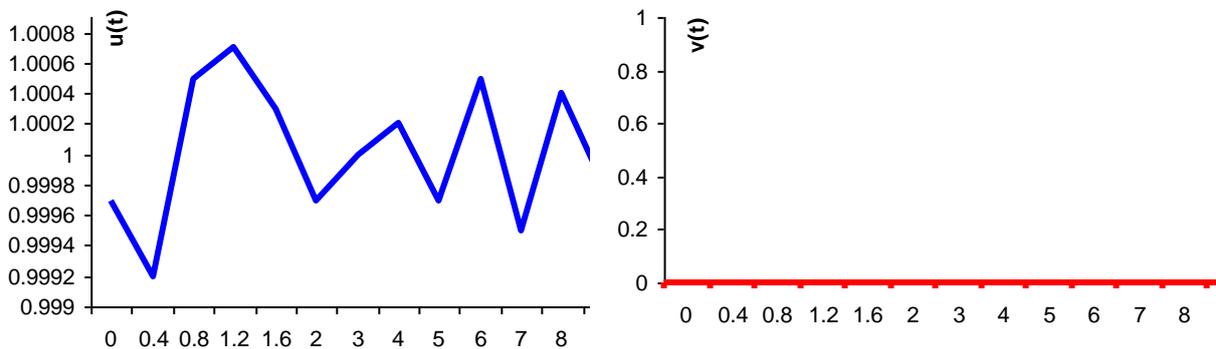
Fig 1: The velocity components for different values of t when M=1; N =1/2; y = 0

We observe that the velocity component v(t) equals zero and the velocity component u(t) approaches to one at y =0 , this means that the fluid is filling the whole space between the two plates .

Case 2.a

Define the variables: $M = 1; \quad N = 1/2; \quad y = 1;$

Next, we obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0.4	0.1525	0.0467
0.8	0.3621	0.0576
1.2	0.4099	0.0628
1.6	0.4251	0.0656
2	0.4302	0.0671
3	0.4303	0.0688
4	0.4291	0.0693
5	0.4285	0.0695
6	0.4275	0.0696
7	0.4283	0.0696
8	0.4272	0.0696
9	0.4281	0.0696

Table (2): The velocity components u and v for different values of t
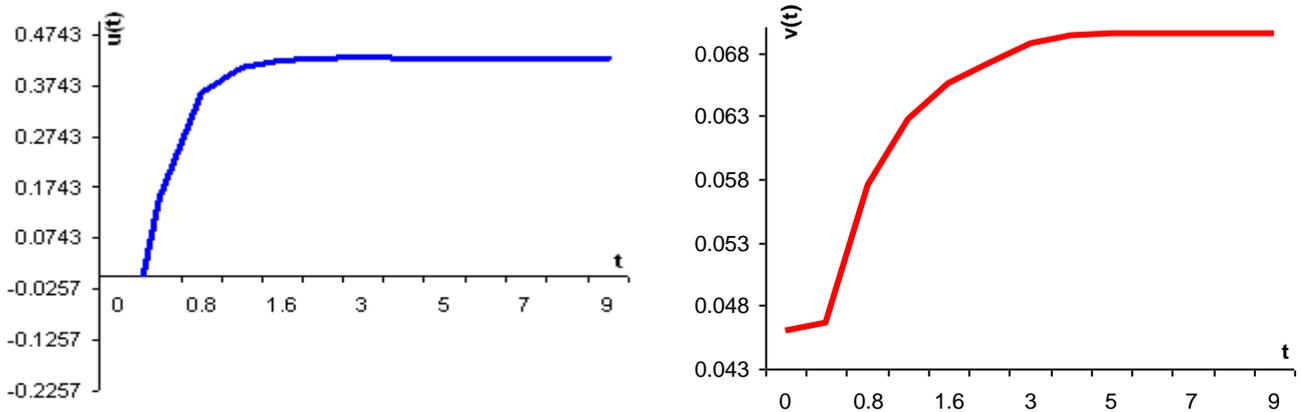
Fig 2: The velocity components for different values of t when M=1; N =1/2; y = 1

Case 3.a

Define the variables:

$M = 1; \quad N = 1 \; ; \; y = 1;$

Next, we obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0.4	0.1517	0.0723
0.8	0.4088	0.0879
1.2	0.4726	0.0964
1.6	0.4927	0.1016
2	0.498	0.1049
3	0.4922	0.1092
4	0.4849	0.1108
5	0.4804	0.1115
6	0.4775	0.1118
7	0.4774	0.1119
8	0.4762	0.1119
9	0.4771	0.1119

Table (3): The velocity components u and v for different values of t
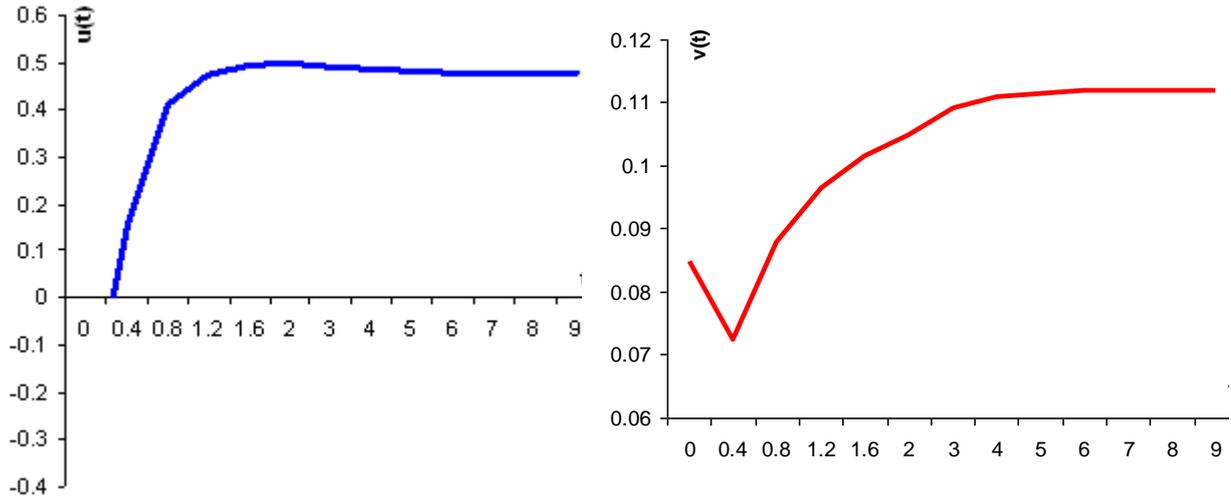
Fig 3: The velocity components for different values of t when M=1; N =1; y = 1

Case 4.a

Define the variables: $M = 2$; $N = 1$; $y = 1$;

Next, we obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0.4	0.1561	0.1008
0.8	0.2027	0.1185
1.2	0.1896	0.1202
1.6	0.1834	0.1201
2	0.1823	0.1199
3	0.1819	0.1199
4	0.1821	0.1199
5	0.1823	0.1199
6	0.1818	0.1199
7	0.1826	0.1199
8	0.1816	0.1199
9	0.1825	0.1199

Table (4): The velocity components u and v for different values of t
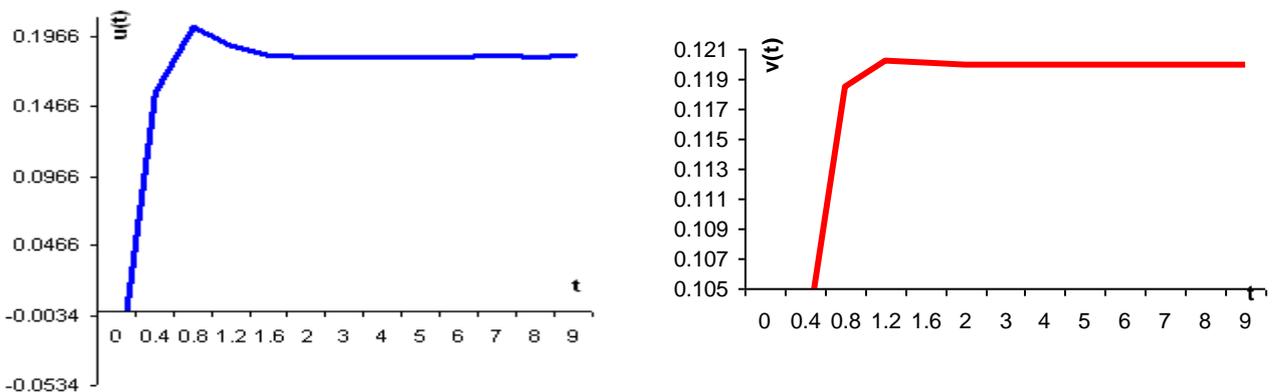
Fig 4: The velocity components for different values of t when M=2; N =1; y = 1

For the next table we'll redefine q as a function of y.

$$q(y, t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \hat{q}(y, s) \, e^{st} \, ds \quad \text{and} \quad \omega := \left(\frac{M^2}{1 + N^2} + s\right) - i\frac{N M^2}{1 + N^2}$$

Case 1.b

Define the variables: $M = 1$; $N = 1/2$; $t = 1/2$;

Next, we define ω and obtain the value η such that $\eta^2 = \omega$ with $\text{Re}(\eta) < 0$.

t	u(t)	v(t)
0	1.9993	0
0.4	0.6713	0.0398
0.8	0.3791	0.0515
1.2	0.1129	0.0458
1.6	-0.1213	0.0301
2	-0.2506	0.0113
3	0.1163	-0.0079
4	-0.1269	0.0017
5	0.0937	0.0007
6	0.1066	-0.0001
7	0.0883	0
8	0.0165	0.0003
9	-0.0636	-0.0001

Table (5): The velocity components u and v for different values of y
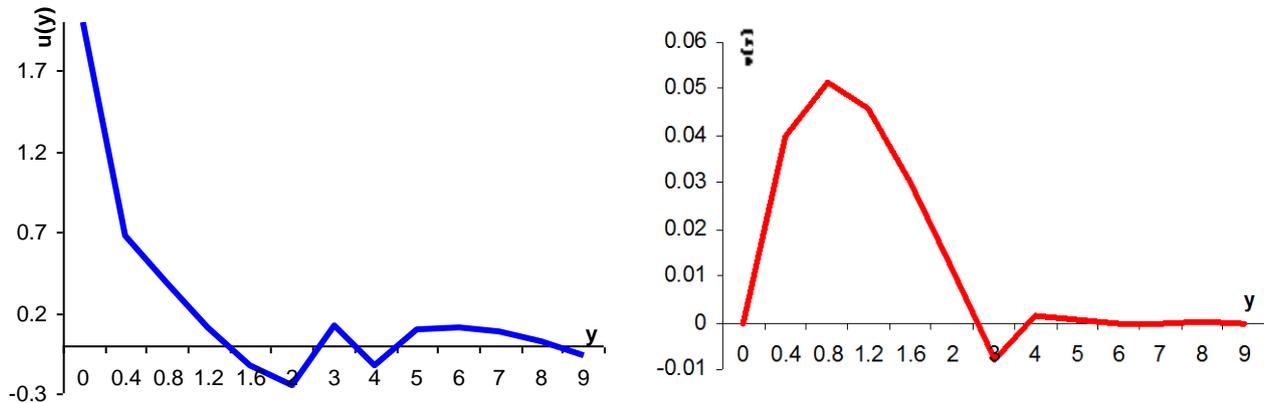
Fig 5: The velocity components for different values of t when M=1;N =1/2; t= 1/2

Case 2.b

Define the variables: M=1;   N=1/2;   t=1;

Next, we define ω and obtain the value η such that $\eta^2 = \omega$ with $\text{Re}(\eta) < 0$.

t	u(t)	v(t)
0	2.0014	0
0.4	0.7076	0.0428
0.8	0.4862	0.059
1.2	0.3116	0.0593
1.6	0.1685	0.0504
2	0.0528	0.0367
3	-0.108	0.0009
4	-0.0053	-0.0091
5	0.0534	0.0034
6	-0.0555	-0.0008
7	0.46	0.0007
8	-0.0224	-0.001
9	-0.0215	0.0007

Table (6): The velocity components u and v for different values of y
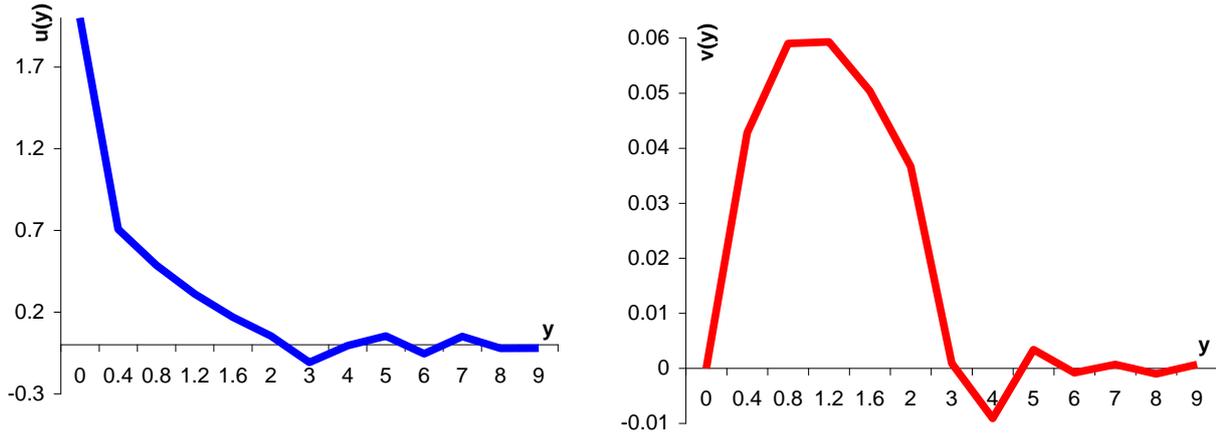
Fig 6: The velocity components for different values of t when M=1; N =1/2; t= 1

Case 3.b

Define the variables: $M = 1;$ $N = 1;$ $t = 1/2;$

Next, we define ω and obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0	1.9993	0
0.4	0.708	0.0558
0.8	0.4111	0.0766
1.2	0.1076	0.0729
1.6	-0.1789	0.053
2	-0.3431	0.0259
3	0.16	-0.0114
4	-0.1691	-0.0003
5	0.1272	0.0009
6	0.1435	0
7	-0.1195	0
8	0.0224	0.0003
9	-0.0859	-0.0001

Table (7): The velocity components u and v for different values of y
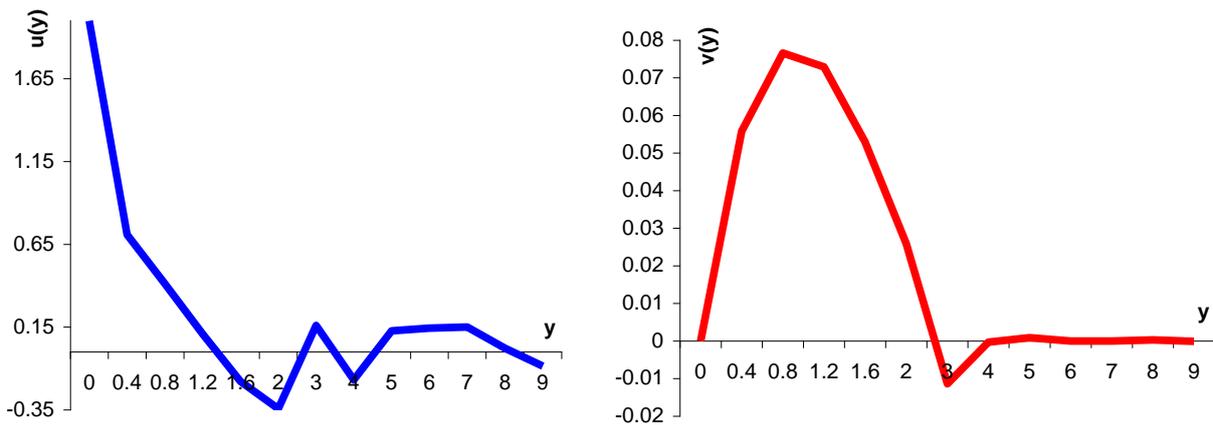
Fig 7: The velocity components for different values of t when M=1;N =1; t= 1/2

Case 4.b

Define the variables:

$M = 3 ;$ $N = 1/2;$ $t = 1/2;$

Next, we define ω and obtain the value η such that $\eta^2 = \omega$ with $\mathrm{Re}(\eta) < 0$.

t	u(t)	v(t)
0	1.9993	0
0.4	0.3273	0.0829
0.8	0.0995	0.054
1.2	0.0301	0.0257
1.6	-0.0064	0.0103
2	-0.0007	0.0031
3	0.0013	-0.0008
4	-0.0005	0.0002
5	0	0
6	0.0004	0
7	0	0
8	0.0002	0
9	-0.0001	0

Table (8): The velocity components u and v for different values of y
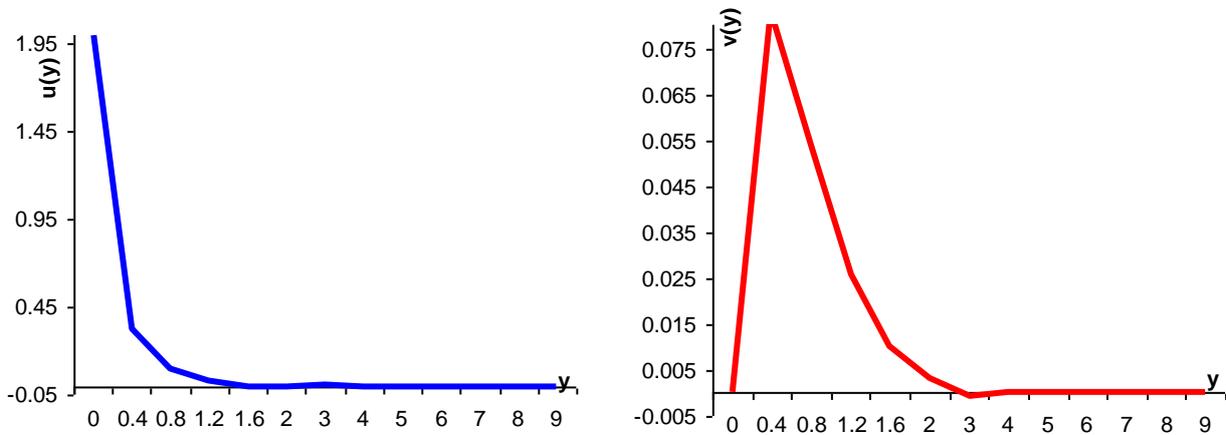
Fig 8: The velocity components for different values of t when M=3;N =1/2; t= 1/2

6. Conclusions

The set of equations describing the MHD flow of an incompressible electrically conducting fluid in the presence of Hall effect are a compensation of the Navier-Stokes equations of fluid dynamics and Maxwell's equations of electromagnetism. These differential equations have been solved analytically and numerically. It must be revealed that only in few special cases an exact solution can be obtained. One of these cases occurs when the compressibility effects of the medium are considered to be negligible. That is, the fluid is taken as in compressible and the other fluid properties such as viscosity, thermal conductivity and electrical conductivity are regarded as constants, for the numerical solution of the partial differential equations the velocity components u and v which are dependent on **M** and **N** have been calculated for different values of times **t** and heights **y.** The numerical results show the following observations:

(i) The velocity component **u** increases with the increase of **N** at equal heights of **y** and   attains a steady state earlier with the increase of **N**.
(ii) The velocity component **u** decreases with the increase of **M**.
(iii) Attaining the steady state is delayed as **N** decreases.
(iv) The velocity component **v** increases with the increase of **M**.
(v)The velocity component **v** increases with the increase of parameter **N.**
(vi) When **y** increases at different values of **t**, **u** decreases, while **v** increases for fixed values of **M** and **N**. Moreover the velocity component **u** gets unstable at different values of **N**, and **y** along with the increase of **t**. However, the velocity component **v** often increases as **t** increases.
(vii) Finally, it can be stated that these new numerical results are in close agreement with the analytical solution that has been obtained by the method of Laplace transform.

References

[1] Kh. Abdul Maleque and Md. Abdus Sattar, The effects of variable properties and Hall current on steady MHD laminar convective fluid flow due to a porous rotating disk, International Journal of Heat and Mass Transfer, 48 (23-24) , 4963 – 4972 (2005) .

[2] C. Chang and J. Yen, Rayleigh's Problem in Magnetohydrodynamics, Phys. Fluids, 2, 239 (1959).

[3] A. Gupta, On the flow of electrically conducting fluid near an accelerated plate in the presence of a magnetic field, J. Phys . Soc. Japan, 15, pp.1894-1897 (1960).

[4] H. Hashimoto, High Temperature Gas Reaction Specimen Chamber for an Electron Microscope, Prog . Theo . Phys. 24, 35 (1962).

[5] M. Kinyanjui , J. Kwanza and S. Uppal, Magnetohydrodynamic 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 , 42 (8) , pp. 917 – 931 (2001).

[6] G. Ludford, On the flow of a conducting fluid past a magnetized sphere, Arch. Rat. Mech . Anal. 3, pp. 102-122 (1959).

[7] H. Mohanty, Hydromagnetic Rayleigh Problem with Hall Effect, Czech . J. Phys. 27, pp.1111-1116 (1977).

[8] R. Nanda and A . Sundaram, Boundary layer growth of an infinite flat plate in magnetohydrodynamics, ZAMP, 13, no.10. pp. 483-489 (1962).

[9] I . Pop, The effect of Hall current on hydromagnetic flow near an accelerated plate, J. Math. Phys. Sci. 5, 293, (1971).

[10] J. Roscizewski , On the flow of electrically conducting fluids over a flat plate in the presence of a transverse magnetic field, Arch . Rat. Mech. Anal.16, 230 (1964).

[11] V . Rossow, On Rayleigh's Problem in Magnetohydrodynamics, Phys. Fluids, 3, 395 (1960).

[12] H. Sato, The Hall effect in the viscous flow of ionized gas between parallel plates under transverse magnetic field, J. Phys . Soc. Japan, 16, pp. 1427-1433 (1961).

[13] H. Schlichting, Boundary- Layer Theory, New York: McGraw – Hill (1955).

[14] A. Sherman and G. Sutton, Engineering Magnetohydrodynamics, New York: McGraw –Hill (1965).

[15] V. Soundalgekar, Mass- transfer effects on the flow past an impulsively started infinite vertical plate with variable temperature or constant heat flux, Astrophysics and Space Science, 100, 1-2, pp.159-164 (1983).

[16] H. Sulieman, The effect of Hall current on Magnetohydrodynamics flow, M.Sc thesis,   Palestine polytechnic University, Hebron (2009).

[17] T. Yaminishi, Preprint, 17th Annual Meeting, Phys. Soc. Japan, Osaka, 5, pp.29 (1962).

XML-based agent communication and migration for distributed applications in Mobile-C

Khaoula ADDAKIRI
Department of Mathematics and Computer Science,
Université Hassan 1er, FSTS, LABO LITEN
Settat, Morocco

Mohamed BAHAJ
Department of Mathematics and Computer Science,
Université Hassan 1er, FSTS, LABO LITEN
Settat, Morocco

Noreddine GHERABI
Department of Mathematics and Computer Science,
Université Hassan 1er, FSTS, LABO LITEN
Settat, Morocco

Abstract
Agent technology is emerging as a powerful approach for the development of complex system. Mobile agent is a program that can migrate as a whole around network and can communicate with its environment and other agent. Among application for mobile agent: manufacturing, electronic commerce, network management, health care, and entertainment. In this paper we present the design and implementation of Mobile-C. The goal of the research is to access an XML file in order to find some information and guaranteed that the data of mobile agent is only accessed by one agent on time by using the synchronization.

Keywords: *Mobile agent; Agent communication; Agent communication; synchronization.*

I. INTRODUCTION

A distributed system is a set of autonomous machines connected through a network and composed of distinct software dedicated to the coordination of system activities, and leverage the availability of several distributed resources to provide better scalability.

Mobile agent is an autonomous software entity responsible for executing a programmatic process, which is able to migrate through a network. An agent migrates in a distributed environment between agencies. When an agent migrates, its execution is suspended at the original agency, the agent is transported to another agency in the distributed environment, and, after being re-instantiated at the new agency, the agent resumes execution.

The majority of mobile agent platforms in use are Java-oriented. Multiple mobile agent platforms supporting Java mobile agent code include Mole [1], Aglets [2], Concordia [3], JADE [4], and D'Agents [5]. Adopting a standard language as the mobile agent code language that provides both high-level and low-level functionalities is a good choice to deal with the diversity of distributed applications.

C/C++is a proper choice for such a mobile agent code language because it's provides powerful functions in terms of memory access. A several number of C/C++programs can be used as mobile agent code. Furthermore, C is a language wish can easily interface with a variety of low-level hardware devices. Ara [6, 7] and TACOMA [8] are two mobile agent platforms supporting C mobile agent code, whereas Ara also supports C++ one. Mobile agent code is compiled as byte code [9] and machine code [10] for execution in Ara and TACOMA, respectively.

Mobile-C [11–14] was originally developed as a stand-alone mobile agent platform to support C/C++ mobile agent code. Mobile-C chose an embeddable C/C++ interpreter— Ch [15–17] to run C/C++ mobile agent source code. The interpretive approach can avoid some potential problems, such as platform portability, secure execution, and system implementation issues that could be induced by the compiling approach. Mobile agent migration in Mobile-C is achieved through Foundation for Intelligent Physical Agents (FIPA) agent communication language (ACL) messages. Using FIPA ACL messages for agent migration in FIPA compliant agent systems simplifies agent platform, reduces development effort and easily achieves inter-platform migration through well-designed communication mechanisms provided in the agent platform. Messages for agent communication and migration are expressed in FIPA ACL and encoded in XML.

In this paper, our approach is to access an XML file in order to search some information. The remainder of the article is structured as follows. Section 2 introduces the architecture of Mobile-C. Section 3 shows two types of messages in Mobile-C, agent communication messages and mobile agent messages and presents how mobile agent migrate from multiple hosts. Section 4 gives an example of mobile agent that access to XML data and illustrates the synchronization support in Mobile-C.

II. THE ARCHITECTURE OF MOBILE-C

The system of mobile-C is shown in figure1. Agencies are the major building blocks and reside in each node of a network system, in wish agents reside end execute. They also serve as "home bases" for locating and messaging agents, migrating mobile agents, collecting knowledge about a group of agents, and providing an environment in which a mobile agents executes. The core of an agency provides local service for agents and proxies remote agencies .An agent platform represents the minimal functionality required by an agencies in order to support the execution of agents. The main of an agency and their functionalities can be summarized as follows [18]:

- Agent Management system (AMS): The AMS manages the life cycle of agents in the system. It controls the creation, registration, retirement, migration and persistence of agents. AMS maintains a directory of agents identifiers (AID).Each agent must register with an AMS in order to get a valid AID.

- Agent Communication Channel (ACC): The ACC routes messages between local and remote entities and realizing using an agent communication language (ACL).

- Agent Security Manager (ASM): The ASM is responsible for maintaining security policies for platform and infrastructure.

- Directory Facilitator: DF serves yellow page services. Agents in the system can register their services with DF for providing to the community. They can also look up required services with DF.

- Agent Execution Engine (AEE): AEE serves as the execution environment for the mobile agents. Mobile agents must reside inside an engine to execute. AEE has to be platform independent in order to support a mobile agent executing in a heterogeneous network.

Figure1.The system architecture of agencies in Mobile-C.

III. MESSAGES AND MIGRATION OF MOBILE AGENT IN MOBILE-C
III.1. Messages in Mobile-C

In Mobile-C there are two types of messages, agent communication messages, and mobile agent messages. A sample agent communication message is from agent-a to agent-b as shown in Figure 2. The message is encoded in XML. In Figure 2, the sender and intended recipient of the message are identified by their agent-identifiers. For the sample message, the sender and receiver agent names are X and Y, respectively. The sender and receiver agent addresses are http://1.fsts.ac.ma:5120 and http://2.fsts.ac.ma: 5120, respectively.

```
<?xml version="1.0" ?>
```

```
<sender>
  <agent_identifier>
    <name>X </name>
    <adresse>
    <url> http://1.fsts.ac.ma:5120</url>
    </adresse>
  </agent_identifier>
</sender>
<receiver>
  <agent_identifier>
    <name>Y</name>
    <adresse>
    <url> http://2.fsts.ac.ma:5120</url>
    </adresse>
  </agent_identifier>
</receiver>
```

Figure2.An ACL message represented in XML

A mobile agent message contains general information about the mobile agent and the tasks performed by the agent in a remote host. The general information of mobile agent includes the name, the owner and the home agent where the mobile agent is created. The task information contains number of tasks, description of tasks and code of each task. During the migration, the task performed by the mobile agent will be encapsulated into agent messages. At the end of the migration, all results of tasks must be sending back to the home agency.

III.2 The migration of mobile agent in Mobile-C

Mobile agent is a software agent who is able to migrate from one host to another in a network and resume the execution in the new host. The migration and the execution of mobile agents are supported by a mobile agent system. In previous studies, Chen et al have developed a mobile agent system called Mobile-C. The Mobile-C supports weak migration. The task of a mobile agent can be divided into several subtasks which can be executed in different hosts and listed in a list of tasks as shown in figure 3. The task list can be modified by adding new subtasks and new conditions. Changing dynamically the task list improves the flexibility of a mobile agent. Thus, once we start the execution of a subtask in a host, the mobile agent cannot move until the end of execution.

Figure3. Agent migration based on a task list and a task progress pointer.

Mobile agent migration is achieved through ACL mobile agent messages encoded to XML, which convey mobile agents as the content of a message. Mobile agent message contains the data state and the code of an agent. The data state of mobile agent include general information about mobile agent as agent name, agent owner and agent home , also the tasks that mobile agent will performed in certain host. The data state and code will be wrapping up into an ACL message and transmitted to a remote host trough Agent Communication Channel. Mobile agent migration based on ACL messages is simple and effective for agent migration in FIPA compliant systems because these systems have mandatory mechanisms for message communication, transmission and procession.

IV. SYNCHRONIZATION SUPPORT IN MOBILE-C

One of the advantages of mobile agents is able to migrate to different hosts to perform tasks based on resources available in remote hosts. The purpose of the mobile agent in this simulation example is to access XML system book and to use a synchronization function as mutex in order to protect a resource that may be shared between two agents. The XML data files store information about the book, the borrower of this book and time of keeping it.

A mobile agent dispatched by an agency in the host fsts1 visits remote host fsts2 and fsts3. Figure 5 shows part of the mobile agent message sent from host fsts1 to host fsts2. The agent transports three kinds of information. First, information about itself including the name, the owner and the home address. Second, global information about the task it has to do. The statement <TASK task="2" num="0"> shows that this mobile agent has two tasks to perform and no task has been finished yet. Third, overall information about the task including the name of the task's return variable, the persistence of the agent, the completeness of

the task, the host to perform the task, and most importantly, the mobile agent code is C/C++ source code that implements the task. Since the persistent is set to 1, the agent will not be removed from an agency once its code has been executed.

```
<NAME> mobagent </NAME>
<OWNER> IEL </OWNER>
<HOME> fsts1fsts.ac.ma:5125 </HOME>
<TASKS task= "2" num= "0">
  <DATA num ="0"
          name = "results_fsts2"
          persistent="1"
          complete = "0"
          server = "fsts2.fsts.ac.ma.5138">
<AGENT_CODE>
Mobile agent code on fsts2
</AGENT_CODE>
</DATA>
  <DATA num ="0"
          name = "results_fsts3"
          persistent="1"
          complete = "0"
          server = " fsts3.fsts.ac.ma.5135">
<AGENT_CODE>
Mobile agent code on fsts3
</AGENT_CODE>
 </DATA>
</TASK>
```

Figure 5: the content of the mobile agent message from the host bird1 to iel2

The task of the mobile agent on fsts2 machine is to access an XML system book information file listed in figure 6, and find the date of borrowing the book and returning it. Function parseNode () is a typical C XML processing program. It can be executed interpretively without the need of compilation in our system. The function searches each node of the XML file and retrieves the date of sortie and return of the book. Also using a variable synchronization in order to guaranteed that data of mobileagent1vis only accessed by one agent on time.

```
<?xml version="1.0" ?>
<!DOCTYPE  SYSTEM BOOK "Book.dtd">
<Book>
<Title>Les réseaux </Title>
<Author> A.Tanebaum </Author>
<Price>250 </Price>
<LoanList>
<Loan>
<borrower>Tarek Amine </borrower>
<Sortie>25/09/2011</Sortie>
<Return>02/11/2011</ Return >
</Loan>
</LoanList>
</Book>
```

Figure 6. The content of an XML system book file.

As shown in Program 1, the mobile agent mobileagent1 performs a ParseNode operation, it's includes locking the

mutex via the function mc_MutexLock() to guaranteed that the desired service and the agent providing the service on the local agency are protected and the simultaneous access of the desired service be avoided .After that , finding the date of borrowing the book by calling parseNode () through the function mc_CallAgentFunc(), and unlocking the mutex via the function mc_MutexUnlock(). .

```
<?xml version="1.0"?>
<MESSAGE message="MOBILE_AGENT">
 <MOBILE_AGENT>
  <AGENT_DATA>
   <NAME>mobileagent1</NAME>
   <OWNER>fsts</OWNER>
   <HOME>fsts1.fsts.ac.ma:5050</HOME>
   <TASKS task="1" num="0">
   <TASK num="0"
       persistent="1"
       name="no-return"
       complete="0"
       server="fsts2.fsts.ac.ma:5130">
   </TASK>
   <AGENT_CODE>
    <![CDATA[
#include <stdlib.h>
#include <string.h>
int main() {
 int i, numService = 1, mutex_id = 55, *agentID,
numResult;
 char *funcname = "ParseNode", **service, **agentName,
**serviceName;
 MCAgent_t agent;
 service = (char **)malloc(sizeof(char *)*numService);
 for(i=0; i<numService; i++) {
  service[i] = (char
*)malloc(sizeof(char)*(strlen(funcname)+1));
 }
 strcpy(service[0], funcname);
 mc_SearchForService(service[0], &agentName,
&serviceName, &agentID, &numResult);
 if(numResults < 1) {
  /* No agent is found to have provided such a service. */
  mc_RegisterService(mc_current_agent, service,
numService);
 }
 else {
  /* An existing agent is found to have provided such a
service. */
  mc_MutexLock(mutex_id);
  mc_DeregisterService(agentID[0], service[0]);
  mc_RegisterService(mc_current_agent, service,
numService);
  mc_MutexUnlock(mutex_id);
  mc_DestroyServiceSearchResult(agentName,
serviceName, agentID, numResult);
 }

 for(i=0; i<numService; i++) {
  free(service[i]);
 }
 free(service);
 return 0;
```

```
}
void ParseNode (xmlDocPtr doc,xmlNodePtr cur) {
static int i;
i++;
while(cur!=NULL);{
  if(cur->type==XML_ELEMENT_NODE){
   if(!(xmlStrcmp(cur-name,(const xmlChar*)"Sortie"))){
      results-iel2[1]=atof(xmlNodeListGestring(doc,
          cur->xmlChildrenNode,1));
      printf(" the date of sortie of the book is
          %f\n",results_fsts2[1];
   }
      parsenode(doc,cur->xmlchildrenNode)
   }
cur = cur->next;
}
i--;
return 0;
}
    ]]>
   </AGENT_CODE>
   </TASKS>
  </AGENT_DATA>
 </MOBILE_AGENT>
</MESSAGE>
```

Program1.A mobile agent that contains a global variable and defines fonctions

Likewise, as shown in Program3, the mobile agent mobileagent3 locks the mutex, finding the date of returning the book by calling parseNode (), and unlocks the mutex afterwards.

```
<?xml version="1.0"?>
<MESSAGE message="MOBILE_AGENT">
 <MOBILE_AGENT>
  <AGENT_DATA>
   <NAME>mobileagent2</NAME>
   <OWNER>fsts</OWNER>
   <HOME>fsts1.fsts.ac.ma:5050</HOME>
   <TASKS task="1" num="0">
   <TASK num="0"
       persistent="1"
       name="no-return"
       complete="0"
       server="fsts2.fsts.ac.ma:5130">
   </TASK>
   <AGENT_CODE>
    <![CDATA[
#include <stdlib.h>
#include <string.h>
int main() {
 int i, numService = 1, mutex_id = 55, *agentID,
numResult;
 char *funcname = "matrix_operate", **service,
**agentName, **serviceName;
 MCAgent_t agent;
 service = (char **)malloc(sizeof(char *)*numService);
 for(i=0; i<numService; i++) {
  service[i] = (char
*)malloc(sizeof(char)*(strlen(funcname)+1));
 }
```

```
strcpy(service[0], funcname);
 mc_SearchForService(service[0], &agentName,
&serviceName, &agentID, &numResult);
 if(numResults < 1) {
   /* No agent is found to have provided such a service. */
   mc_RegisterService(mc_current_agent, service,
numService);
 }
 else {
   /* An existing agent is found to have provided such a
service. */
   mc_MutexLock(mutex_id);
   mc_DeregisterService(agentID[0], service[0]);
   mc_RegisterService(mc_current_agent, service,
numService);
   mc_MutexUnlock(mutex_id);
   mc_DestroyServiceSearchResult(agentName,
serviceName, agentID, numResult);
 }

 for(i=0; i<numService; i++) {
  free(service[i]);
 }
 free(service);
 return 0;
}
void ParseNode (xmlDocPtr doc,xmlNodePtr cur) {
static int i;
i++;
while(cur!=NULL);{
  if(cur->type==XML_ELEMENT_NODE){
   if(!(xmlStrcmp(cur-name,(const xmlChar*)"return"))){
      results-fsts2[1]=atof(xmlNodeListGestring(doc,
         cur->xmlChildrenNode,1));
      printf("The date of retour of the book is
         %f\n", results_fsts2 [1];
      }
      parseNode(doc,cur->  xmlchildrenNode)
      }
cur = cur->next;
}
i--;
return 0;
}
   ]]>
   </AGENT_CODE>
  </TASKS>
  </AGENT_DATA>
 </MOBILE_AGENT>
</MESSAGE>
```

Program2.A code of mobile agent performing the second task in the fsts2

After visiting the host fsts2 , the mobile agents visit the host fsts3 .Likewise, as shown in Program 1 and 2, the task of the mobile agent on fsts3 is locks the mutex, reads the xml file, and unlocks the mutex afterwards.

V.   CONCLUSION

This article presents an XML-based approach for agent communication, and migration in mobile-C. Mobile-C conforms to the IEEE FIPA standards, it's integrates an embeddable C/C++ interpreter into the platform as a mobile agent execution engine in order to support mobile agents. Mobile agents, including its data state and code, are carries to a remote agent platform via ACL messages wich will be encoded in XML, and the execution of mobile agents is resumed by a task progress pointer. Our work shows that using XML to encode different types of messages is simple, and easy to change .Thus, the synchronization functions guaranteed the protection of shared resources by using the mutex in multiplt hosts.

REFERENCES

[1] J. Baumann, F. Hohl, K. Rothermel, M. Strasser, W. M .Theilmann: A mobile agent system. Software—Practice and Experience 2002; 32(6):575–603.

[2] D. Lange, M.Oshima. Programming and Deploying Java Mobile Agents with Aglets. Addison-Wesley: MA, 1998.

[3] D.Wong, N.Paciorek, T.Walsh, J.DiCelie, M.Young, B.Peet. Concordia: An infrastructure for collaborating mobile agents. Proceedings of the First International Workshop on Mobile Agents (MA'97) (Lecture Notes in Computer Science, vol. 1219). Springer: Berlin, 1997; 86–97.

[4] F.Bellifemine, G.Caire, A.Poggi, G.Rimassa.JADE: A software framework for developing multi-agent applications.Lessons learned. Information and Software Technology 2008; 50(1–2):10–21.

[5] R.Gray, G.Cybenko, D.Kotz,R.Peterson, D.Rus. D'Agents: Applications and performance of a mobile-agent system. Software—Practice and Experience 2002; 32(6):543–573.

[6] H.Peine. Run-time support for mobile code. PhD Dissertation, Department of Computer Science, University of Kaiserslautern, Germany, 2002.

[7] H.Peine .Application and programming experience with the Ara mobile agent system. Software—Practice and Experience 2002; 32(6):515–541.

[8] D.ohnansen, K.Lauvset, R.V.Renesse, F.B. Schneider, N.P. Sudmann, K. Jacobsen. A TACOMA retrospective.Software—Practice and Experience 2002; 32(6):605–619.

[9] MACE—Mobile agent code environment. Available at: http://wwwagss.informatik.uni-kl.de/Projekte/Ara/mace.html [last modified 10 August 2004].

[10] N.P.Sudmann,D.Johansen. Adding mobility to non-mobile web robots. Proceedings of the IEEE ICDCS00 Workshop on Knowledge Discovery and Data Mining in the World-wide Web, Taipei, Taiwan, 2000; 73–79.

[11] B.Chen, H.H.Cheng. A run-time support environment for mobile agents. Proceedings of ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications, No. DETC2005-85389, Long Beach, CA, September 2005.

[12] B.Chen,H.H.Cheng,J.Palen. Mobile-C: A mobile agent platform for mobile C/C++ agents. Software—Practice and Experience 2006; 36(15):1711–1733.

[13] B.Chen, D.Linz, H.H.Cheng. XML-based agent communication, migration and computation in mobile agent systems. Journal of Systems and Software 2008; 81(8):1364–1376.

[14] Mobile-C: A multi-agent platform for mobile C/C++ code. Available at: http://www.mobilec.org [last modified 12 May 2009].

[15] H.H.Cheng. Scientific computing in the Ch programming language. Scientific Programming 1993; 2(3):49–75.

[16] H.H.Cheng.Ch: A C/C++ interpreter for script computing. C/C++ User's Journal 2006; 24(1):6–12.

[17] Ch—An embeddable C/C++ interpreter. Available at: http://www.softintegration.com [last modified 15 April 2009].

[18] B.Chen, H.H.Cheng, J.Palen. Integrating mobile agent technology with multi-agent systems for distributed traffic detection and management systems, Transportation Research Part C 17 (2009) 1–10.

MEDICAL IMAGE SEGMENTATION USING FUZZY C-MEANS CLUSTERING AND MARKER CONTROLLED WATERSHED ALGORITHM

M.C.Jobin Christ[1], Dr.R.M.S.Parvathi[2]

[1]Adhiyamaan College of Engineering, Dr.MGR Nagar, Hosur, India
[2]Sengunthar College of Engineering, Tiruchengode, India

Abstract

Segmentation plays a vital role in medical imaging. Segmentation of an image is the partition or separation of the image into disjoint regions of similar feature. We propose a method that integrates Fuzzy C-Means (FCM) clustering and marker controlled watershed segmentation algorithm for medical image segmentation. The use of the usual watershed algorithm for medical image analysis is common because of its advantages, such as always being able to construct a complete division of the image. However, its downsides include over-segmentation and sensitivity to false edges. We concentrate on the downsides of the usual watershed algorithm when it is applied to medical images by using Fuzzy C-Means clustering to produce a primary segmentation of the image before we apply marker controlled watershed segmentation algorithm to it. The Fuzzy C-Means clustering is an unsupervised learning algorithm, while the marker controlled watershed segmentation algorithm makes use of automated thresholding on the gradient magnitude map on the initial partitions to reduce the number of false edges and over-segmentation. The proposed algorithm is compared with conservative watershed method.

Keywords: Clustering, FCM, Watershed, Segmentation, Medical Image.

1. Introduction

Image segmentation is a vital method for most medical image analysis tasks. Having good segmentations will help clinicians and patients as they provide vital information for 3-D visualization, surgical planning and early disease recognition. Image segmentation algorithms are classified into two types, supervised and unsupervised. Unsupervised algorithms are fully automatic and partition the regions in feature space with high density. The different unsupervised algorithms are Feature-Space Based Techniques, Clustering (K-means algorithm, C-means algorithm, E-means algorithm), Histogram thresholding, Image-Domain or Region Based Techniques (Split-and-merge techniques, Region growing techniques, Neural-network based techniques, Edge Detection Technique), Fuzzy Techniques, etc. The watershed segmentation technique has been widely used in medical image segmentation. Watershed transform is used to segment gray matter, white matter and cerebrospinal fluid from magnetic resonance (MR) brain images. The method originated from mathematical morphology that deals with the topographic representation of an image. Watersheds are one of the typical regions in the field of topography. A drop of the water falling it flows down until it reaches the bottom of the region. Monochrome image is considered to be a height surface in which high-altitude pixels correspond to ridges and low-altitude pixels correspond to valleys. This suggestion says if we have a minima point, by falling water, region and the frontier can be achieved. Watershed uses image gradient to initial point and region can get by region growing. The accretion of water in the neighborhood of local minima is called a catchment basin. Watershed refers to a ridge that divides areas shattered by different river systems. A catchment basin is the environmental area draining into a river or reservoir. If we consider that bright areas are high and dark areas are low, then it might look like the plane. With planes, it is natural to think in terms of catchment basins and watershed lines. Two approaches are there to find watershed of an image,

[1] Rainfall approach
[2] Flooding approach

In rainfall approach, local minima are found all through the image, and each local minima is assigned an exclusive tag. An intangible water drop is placed at each untagged pixel. The drop moves to low amplitude neighbor until it reaches a tagged pixel and it assumes tag value. In flooding approach, intangible pixel holes are pierced at each local minima. The water enters the holes and takings to fill each catchment basin. If the basin is about to overflow, a dam is built on its neighboring ridge line to the height of high altitude ridge point. These dam borders correspond to the watershed lines. Advantages of the watershed transform include the fact that it is a fast, simple and intuitive method. More importantly, it is able to produce a entire division of the image in separated regions even if the contrast is poor, thus there is no need to carry out any post processing work, such as contour joining. Its limitations will include over-segmentation and sensitivity to noise. There has also

been an increasing interest in applying soft segmentation algorithms, where a pixel may be classified partially into multiple classes, for MR brain images segmentation. Clustering is a method of grouping a set of patterns into a number of clusters such that similar patterns are assigned to one cluster. Each pattern can be represented by a vector having many attributes. Clustering technique is based on the computation of a measure of similarity or distance between the respective patterns. The Fuzzy C-means clustering algorithm is a soft segmentation method that has been used extensively for segmentation of MR brain images. In this work, we use Fuzzy C-means clustering to produce a primary segmentation of the image before we apply the marker controlled watershed segmentation algorithm.

2.Fuzzy C-Means algorithm

Fuzzy C-Means clustering (FCM), also called as ISODATA, is a data clustering method in which each data point belongs to a cluster to a degree specified by a membership value. FCM is used in many applications like pattern recognition, classification, image segmentation, etc. FCM divides a collection of n vectors c fuzzy groups, and finds a cluster center in each group such that a cost function of dissimilarity measure is minimized. FCM uses fuzzy partitioning such that a given data point can belong to several groups with the degree of belongingness specified by membership values between 0 and

1. This algorithm is simply an iterated procedure. The algorithm is given below.

1) Initialize the membership matrix U with random values between 0 and 1.
2) Calculates c fuzzy cluster center c_i, $i = 1,\ldots c$., using the following equation,

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}$$

3) Compute the cost by the following equation. Stop if either it is below a certain threshold value or its improvement over previous iteration.

$$J(U, c_1,\ldots, c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j}^{n} u_{ij}^{m} d_{ij}^{2}$$

4) Compute a new U by the equation. Go to step 2.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}}$$

There is no guarantee ensures that FCM converges to an optimum solution. The performance is based on the initial cluster centers. FCM also suffers from the presence of outliers and noise and it is difficult to identify the initial partitions.

3.Marker Controlled Watershed Segmentation Algorithm

Segmentation using the watershed transforms works well if you can identify, or mark, foreground objects and background locations. The gradient magnitude of the primary segmentation is obtained by applying the Sobel operator. The Canny edge detector was also experimented on, but it was found that the results obtained by both methods are comparable. Hence, we decided on the Sobel filter as the Canny edge detector has higher complexity. In addition, the Sobel filter has the advantage of providing both a differencing and smoothing effect. Marker-controlled watershed segmentation follows this basic procedure:

1. Compute a segmentation function. This is an image whose dark regions are the objects we are trying to segment.
2. Compute foreground markers. These are connected blobs of pixels within each of the objects.
3. Compute background markers. These are pixels that are not part of any object.
4. Modify the segmentation function so that it only has minima at the foreground and background marker locations.
5. Compute the watershed transform of the modified segmentation function.

4.Proposed Methodology

The proposed methodology is a two stage process. The first process uses K-means clustering to produce a primary segmentation of the input image, while the second process applies the marker controlled watershed segmentation algorithm to the primary segmentation to obtain the final segmentation map. Fig.1 describes the flowchart proposed method.

5.Results

We applied our proposed methodology of Fuzzy C-Means clustering integrated with marker controlled watershed algorithm to MR brain images of the head and obtained general segmentation maps of them. We evaluated the performance of our proposed methodology by comparing it with conservative watershed algorithm. The use of Fuzzy C-means clustering before applying marker controlled watershed segmentation algorithm has achieved the objective of reducing the problem of over-segmentation when applied to MR brain images. The segmentation results are displayed in Fig.2.
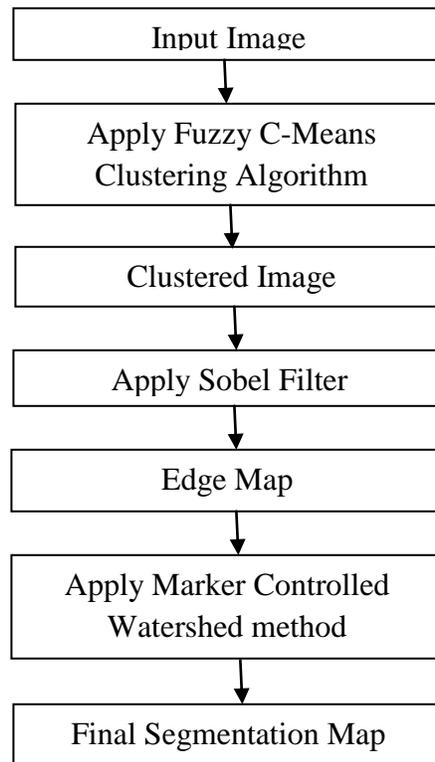
Fig.1 Flow diagram of proposed method

Input Image	Conservative Watershed Algorithm	After Fuzzy C-Means Clustering	Final Segmentation
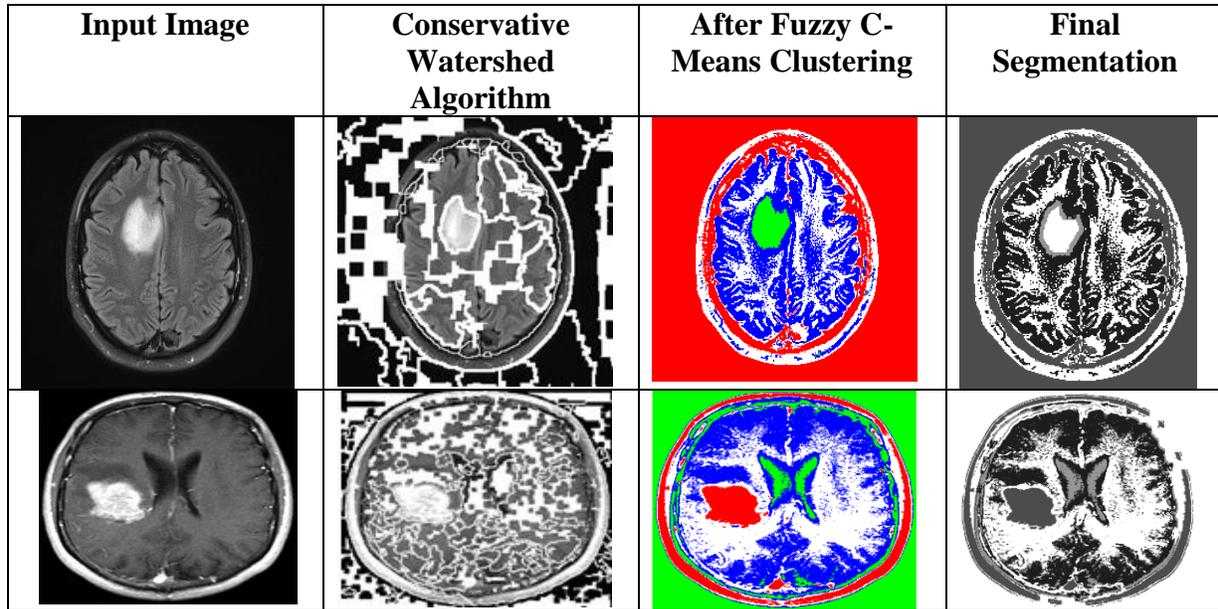
Fig 2. Comparing segmentation maps obtained using proposed method against conservative

6.Discussion

A method which integrated the Fuzzy C-Means clustering algorithm with the marker controlled watershed segmentation algorithm has been proposed. It addressed the limitations of the conservative watershed algorithm, which included over segmentation. The experimental results had shown that our proposed method of using Fuzzy C-means clustering to obtain a primary segmentation of MR brain images before applying the marker controlled watershed segmentation to them is effective. By reducing the amount of over segmentation, we obtained a segmentation map which is more diplomats of the several anatomies in the medical images.

7.References

[1] K.Hari Krishna, R.V.S.Sathyanarayana,"Medical Image Segmentation using Marker Controlled Watershed Algorithm", IJEEE, Vol.1(1), 2009, pp 92-96.

[2] V. Grau, R. Kikinis, M. Alcaniz, S.K. Warfield, "Cortical gray matter segmentation using an improved watershedtransform", Proceedings 25th Annual Int. Conf. of Engineering in Medicine and Biology Society, Vol.1, 2003, pp. 618-621

[3] V. Grau, A.U.J. Mewes, M. Alcaniz, R. Kikinis, S.K. Warfield, "Improved watershed transform for medical image segmentation using prior information", IEEE Transactions on Medical Imaging, Vol.23(4), 2004, pp. 447-458

[4] L. Vincent, P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13(6), pp. 583-598

[5] J.B.T.M. Roerdink, A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies", Fundamental Informaticae, Vol.41, 2000, pp. 187-228

[6] J.C. Bezdek, L.O. Hall, L.P. Clarke, "Review of MR image segmentation techniques using pattern recognition", Medical Physics, Vol. 20(4), 1993, pp. 1033-1048

[7] Veronica S. Moertini, "Introduction to five data clustering algorithms", INTEGRAL, 2002, Vol. 7, No. 2.

[8] H.P. Ng, S.H. Ong, K.W.C. Foong, W.L. Nowinski, "An improved watershed algorithm for medical image segmentation", Proceedings 12th International Conference on Biomedical Engineering, 2005

[9] C.W. Chen, J. Luo, K.J. Parker, "Image segmentation via adaptive K-mean clustering and knowledge based morphological operations with biomedical applications", IEEE Transactions on Image Processing, Vol.7 (12), 1998, pp 1673-1683

[10] R. C. Gonzalez, R. E. Woods, "Digital Image Processing", Addison-Wesley, 1992, pp.419

INTRUSION RESPONSE SYSTEM TO AVOID ANOMALOUS REQUEST IN RDBMS

Akila.L[1], Mrs.DeviSelvam[2]

[1]II M.E CSE,Sri shakthi Institute Of Engineering and Technology,Anna University,Coimbatore
[2]Asst.prof CSE,Sri shakthi Institute Of Engineering and Technology,Anna University,Coimbatore

Abstract:

The intrusion response component of an overall intrusion detection system is responsible for issuing a suitable response to an anomalous request. In the existing system, Intrusion Detection mechanism consists of two main elements, specifically tailored to a DBMS: anomaly detection (AD) system and an anomaly response system. In anomaly response system conservative actions, fine-grained actions, and aggressive actions methods are used. The proposed system mainly concentrates on response policies by using policy matching and policy administration. For the policy matching problem, we propose two algorithms that efficiently search the policy database for policies that match an anomalous request. We also extend the PostgreSQL DBMS with our policy matching mechanism, and report experimental results. The other issue that we address is that of administration of response policies to prevent malicious modifications to policy objects from legitimate users. We propose a novel Joint Threshold Administration Model (JTAM) that is based on the principle of separation of duty. The key idea in JTAM is that a policy object is jointly administered by at least k database administrator (DBAs), that is, any modification made to a policy object will be invalid unless it has been authorized by at least k DBAs out of L. We present design details of JTAM which is based on a cryptographic threshold signature scheme, and show how JTAM prevents malicious modifications to policy objects from authorized users.

Index Terms: PostgreSQL DBMS, Joint Threshold Administration Model, Threshold signatures.

I. INTRODUCTION

Our approach to an ID mechanism consists of two main elements, specifically tailored to a DBMS: an anomalydetection (AD) system and an anomaly response system.The first element is based on the construction of databaseaccess profiles of roles and users, and on the use of such profiles for theADtask.Auser-request that does not conform to the normal access profiles is characterized as anomalous. Profiles can record information of different levels of details;we refer the reader to for additional information and experimental results. The second element of our approach—the focus of this paper—is in charge of taking some actions once an anomaly is detected. There are three main types of response actions that we refer to, respectively, as conservative actions, fine-grained actions, and aggressive actions. The conservative actions, such as sending an alert, allow the anomalous request to go through, whereas the aggressive actions can effectively block the anomalous request. Fine-grained response actions, on the other hand, are neither conservative nor aggressive. Such actions may suspend or taint an anomalous request. A suspended request is simply put on hold, until some specific actions are executed by the user, such as the execution of further authentication steps. A tainted request is marked as a potential suspicious request resulting in further monitoring of the user and possibly in the suspension or dropping of subsequent requests by the same user.

The two main issues that we address in the context of such response policies are that of policy matching and policy administration. Policy matching is the problem of searching for policies applicable to an anomalous request. When an anomaly is detected, the response system must search through the policy database and find policies that match the anomaly. Our ID mechanism is a real-time intrusion detection and response system; thus efficiency of the policy search procedure is crucial. In Section 4, we present two efficient algorithms that take as input the anomalous request details[4], and search through the policy database to find the matching policies. We implement our policy matching scheme in the PostgreSQL DBMS [7], and discuss relevant implementation issues. We also report experimental results that show that our techniques are very efficient.

II.AN OVERVIEW OF RELATED WORK

Administration model is based on the well known security principle of separation of duties (SoD). SoD is a principle whereby multiple users are required in order to complete a given task. As a security principle, the primary objective of SoD is prevention of fraud (insider threats), and user generated errors. Such objective is traditionally achieved by dividing the task and its associated privileges among multiple users.

However, the approach of using privilege dissemination is not applicable to our case as we assume the DBAs to possess all possible privileges in the DBMS. Our approach instead applies the technique of threshold cryptography signatures to achieve SoD. A DBA authorizes a policy operation, such as create or drop, by submitting a signature share on the policy. At least k signature shares are required to form a valid final signature on a policy, where k is a threshold parameter defined for each policy at the time of policy creation. The final signature is then validated either periodically or upon policy usage to detect any malicious modifications to the policies.

The key idea in our approach is that a policy operation is invalid unless it has been authorized by at least k DBAs. We thus refer to our administration model as the Joint Threshold Administration Model (JTAM) for managing response policy objects. To the best of our knowledge, ours is the only work proposing such administration model in the context of management of DBMS objects.

The three main advantages of JTAM are as follows: First, it requires no changes to the existing access control mechanisms of a DBMS for achieving SoD. Second, the final signature on a policy is nonrepudiable, thus making the DBAs accountable for authorizing a policy operation. Third, and probably the most important, JTAM allows an organization to utilize existing man-power resources to address the problem of insider threats since it is no longer required to employ additional users as policy administrators.

III. CREATION OF INTRUSION RESPONSE SYSTEM

The main contributions of this paper can be summarized as follows:
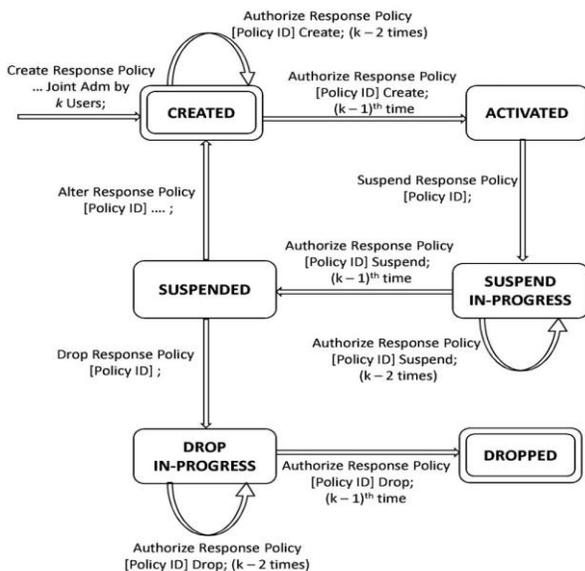
1. We present a framework for specifying intrusion response policies in the context of a DBMS.
2. We present a novel administration model called JTAM for administration of response policies.
3. We present algorithms to efficiently search the policy database for policies that match an anomalous request.
4. We extend the PostgreSQL DBMS with our response policy mechanism, and conduct an experimental evaluation of our techniques.

In this section, we describe the signature share generation, the signature share combining, and the final signature verification operations, in the context of the administrative lifecycle of a response policy object. The steps in the lifecycle of a policy object are policy creation, activation, suspension, alteration, and deletion. The lifecycle is shown in Fig. 1 using a policy state transition diagram. The initial state of a policy object after policy creation is CREATED. After the policy has been authorized by k _ 1 administrators, the policy state is changed to ACTIVATED. A policy in an ACTIVATED state is operational, that is, it is considered by the policy matching procedure in its search for matching policies. If a policy needs to be altered, dropped or made nonoperational, it must be moved to the SUSPENDED state. The transition from the ACTIVATED state to the SUSPENDED state must also be authorized by k _ 1 administrators, before which the policy is in the SUSPEND IN-PROGRESS state. Note that a policy in the SUSPEND IN-PROGRESS state is also considered to be operational. From the SUSPENDED state, a policy can be either moved back to the CREATED state or it can be moved to the DROPPED state. A single administrator can move a policy to the CREATED state from the SUSPENDED state, while a policy drop operation must be authorized by k _ 1 administrators (before which the policy is in the DROP IN-PROGRESS state). We begin our detailed discussion of a policy object's lifecycle with the policy creation procedure.

Fig .1. Policy state transition diagram.

OVERALL PROCESS OF INTRUSION RESPONSE SYSTEM
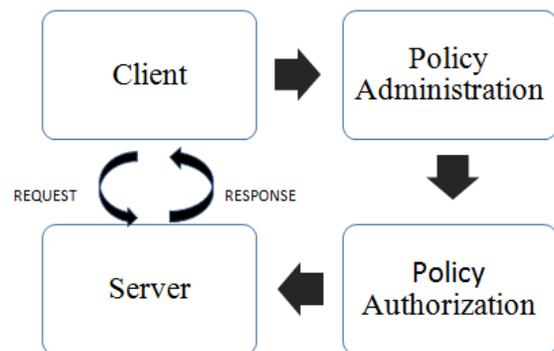
Fig .2.Flow of process

Client requests to the server for performing their operations in the database. For that, the server does the following policy methods like policy administration, policy authorization to find whether the client is intruder or not. After performing the policy methods, the server responds to the client. By the way the intruders are found. The working principle of each components in the fig.2 are explained below,

A.SERVER

We address in the context of such response policies are that of policy matching and policy Administration when an anomaly is detected, the response system must search through the policy database and find policies that match the anomaly. Our ID mechanism is a real-time intrusion detection and response system; thus efficiency of the policy search procedure is crucial.

The second issue that we address is that of administration of response policies. Intuitively, a response policy can be considered as a regular database object such as a table or a view. Privileges, such as create policy and drop policy that are specific to a policy object type can be defined to administer policies. However, a response policy object presents a different set of challenges than other database object types.

Interactive response policy language makes it very easy for the database administrators to specify appropriate response actions for different circumstances depending upon the nature of the anomalous request. The two main issues that we address in context of such response policies are that of policy matching, and policy administration. An anomaly detection (AD) system and an anomaly response system. The first element is based on the construction of database access profiles of roles and users, and on the use of such profiles for the AD task. A user request that does not conform to the normal access profiles is characterized as anomalous. Profiles can record information of different levels of details; we refer the reader to for additional information and experimental results.

B.POLICY ADMINISTRATION

An administration model referred to as the JTAM. The threat scenario that we assume is that a DBA has all the privileges in the DBMS, and thus it is able to execute arbitrary SQL insert, update, and delete commands to make malicious modifications to the policies. Such actions are possible even if the policies are stored in the system catalogs.3 JTAM protects a response policy against malicious modifications by maintaining a digital signature on the policy definition. The signature is then validated either periodically or upon policy usage to verify the integrity of the policy definition.

JTAM is that we do not assume the DBMS to be in possession of a secret key for verifying the integrity of policies. If the DBMS had possessed such key, it could simply create a HMAC (Hashed Message Authentication Code) of each policy using its secret key, and later use the same key to verify the integrity of the policy.

C.POLICY AUTHORIZATION

The detection of an anomaly by the detection engine can be considered as a system event. The attributes of the anomaly, such as user, role, SQL command, then correspond to the environment surrounding such an event. Intuitively, a policy can be specified taking into account the anomaly attributes to guide the response engine in taking a suitable action. Keeping this in mind, we propose an Event- Condition-Action (ECA) language for specifying response policies[1].

A DBA authorizes a policy operation, such as create or drop, by submitting a signature share on the policy. At least k signature shares are required to form a valid final signature on a policy, where k is a threshold parameter defined for each policy at the time of policy creation.

The final signature is then validated either periodically or upon policy usage to detect any malicious modifications to the policies. The key idea in our approach is that a policy operation is invalid unless it has been authorized by at least k DBAs. We thus refer to our administration model as the Joint Threshold Administration Model (JTAM) for managing response policy objects[3].

It requires no changes to the existing access control mechanisms of a DBMS for achieving SoD. Second, the final signature on a policy is non reputable, thus making the DBAs accountable for authorizing a policy operation. Third, and probably the most important, JTAM allows an organization to utilize existing man-power resources to address the problem of insider threats since it is no longer required to employ additional users as policy administrators.

Once a database request has been flagged off as anomalous, an action is executed by the response system to address the anomaly. The response action to be executed is specified as part of a response policy. Such actions may log the anomaly details or send an alert, but they do not proactively prevent an intrusion. Aggressive actions, on the other hand, are high severity responses. Such actions are capable of preventing an intrusion proactively by dropping the request, disconnecting the user or revoking/denying the necessary privileges.

Signature

We describe the signature share generation, the signature share combining, and the final signature verification operations, in the context of the administrative

lifecycle of a response policy object. The steps in the lifecycle of a policy object are policy creation, activation, suspension, alteration, and deletion[8].

It is possible for a malicious administrator to replace a valid signature share with some other signature share that is generated on a different policy definition. However, such attack will fail as the final signature that is produced by the signature share combining algorithm will not be valid. Note that by submitting an invalid signature share, a malicious administrator can block the creation of a valid policy. We do not see this as a major problem since the threat scenario that we address is malicious modifications to existing policies, and not generation of policies themselves.

D.CLIENT

Often clients and servers communicate over a computer network on separate hardware, but both client and server may reside in the same system. A server machine is a host that is running one or more server programs which share their resources with clients. A client does not share any of its resources, but requests a server's content or service function. Clients therefore initiate communication sessions with servers which await incoming requests.Note that implementing the confirmation actions such as a re authentication or a second factor of authentication require changes to the communication protocol between the database client and the server. The scenarios in which such confirmation actions may be useful are when a malicious subject (user/process) is able to bypass the initial authentication mechanism of the DBMS due to software vulnerabilities (such as buffer overflow) or due to social engineering attacks (such as using someone else's unlocked unattended terminal).

Interactive response with the user is not required; the confirmation/resolution/failure actions may be omitted from the policy.

IV. PERFORMANCE EVALUATION

We perform three sets of experiments. The first two experiments report and compare the overhead of the policy matching algorithms. The third experiment reportsresults on the overhead of the signature verification mechanism in JTAM.

In the first experiment, the anomaly assessment is set such that the number of matching policies for an anomaly is kept constant at four. The number of predicates, and correspondingly the number of policies, are varied in order to assess the policy matching overhead time. Fig. 1 shows the policy matching overhead for the two algorithms as a function of the number of predicates. Fig. 2 reports the number of predicates skipped as a function of the number of predicates. As expected, the policy matching overhead time increases linearly with the increase in the number of predicates in the policy database.

Interestingly, the number of predicates skipped in both the algorithms is almost same.

Thus, counter-intuitively, the ordered policy matching algorithm does not lead to a decrease in the number of predicate evaluations. In fact, for larger number of predicates, the policy matching overhead of the ordered predicate algorithm is higher than that of the base policy matching algorithm.
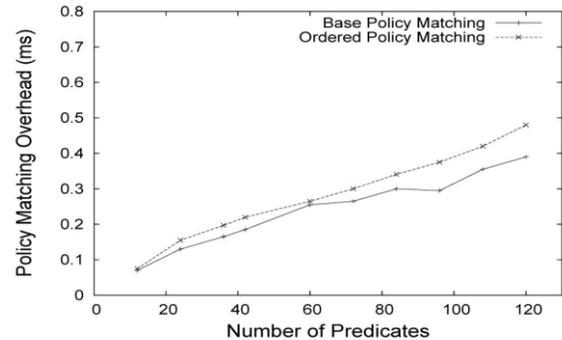
Fig 1. Experiment 1: Number of predicates versus policy matching overhead.

Such increase in matching overhead may be explained by the fact that the predicates evaluated by the ordered policy matching are more computationally expensive than the ones evaluated by the base policy matching algorithm. The key observation from this experiment, however, is that predicate ordering based on the policy-count parameter has no benefits in terms of decreasing the overhead of the policy matching procedure.
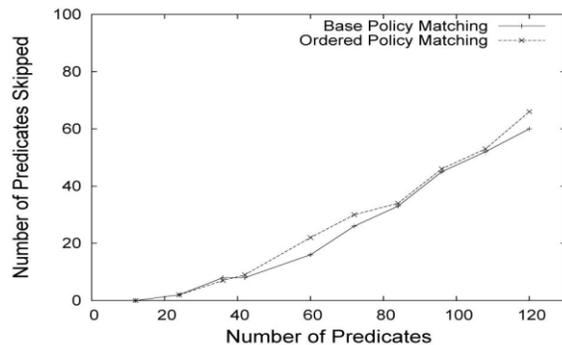
Fig. 2. Experiment 1: Number of predicates versus number of predicates skipped.

In the second experiment, we keep the number of predicates in the policy database constant at 60. The number of policies is also kept constant at 20. The number of matching policies is varied in order to assess the policy matching overhead. Fig. 3 shows the policy matching overhead for the two algorithms as a function of the number of matching policies. As expected, the policymatching overhead increases with the increase in the number of matching policies.
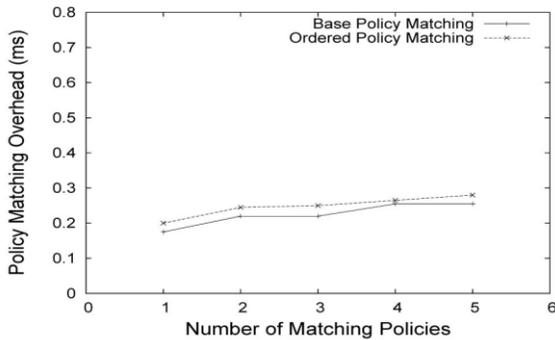
Fig 3. Experiment 2: Number of matching policies versus policy matching overhead

Moreover, in this experiment as well, the overhead of the ordered policy matching algorithm is higher than that of the base policy matching algorithm.
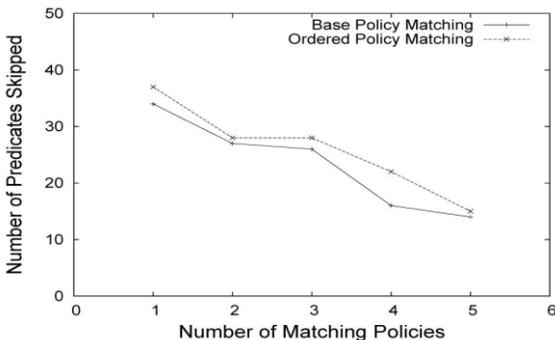
Fig. 4 Experiment 2: Number of matching policies versus number of predicates skipped

Fig. 4 reports the variation in the number of predicates skipped by varying the number of matching policies. For both the algorithms, the number of predicates skipped by the search procedure decreases for increasing numbers of matching policies. Such result is expected since an increase in the number of matching policies leads to an increasing number of predicate evaluations.
Overall, the fist two experiments confirm the low
overhead associated with our policy matching algorithms.

They also show that predicate ordering based on the descending policy-count parameter has no significant impact on reducing the overhead of the policy matching procedure.
Therefore, a better strategy is to create a dedicated DBMS process that periodically polls the policy tables, and verifies the signature on all the policies.

V. CONCLUSION

In this paper, we have described the response component of our intrusion detection system for a DBMS. We presented an interactive Event-Condition-Action type response policy language that makes it very easy for the database security administrator to specify appropriate response actions for different circumstances depending upon the nature of the anomalous request. The two main issues that we addressed in the context of such response policies are policy matching, and policy administration. Specifically, we added support for new system catalogs to hold policy related data, implemented new SQL commands for the policy administration tasks, and integrated the policy matching code with the query processing subsystem of PostgreSQL. The other issue that we addressed is the administration of response policies to prevent malicious modifications to policy objects from legitimate users. We proposed a JTAM, a novel administration model, based on Shoup's threshold cryptographic signature scheme we are currently in the process of implementing the intrusion detection algorithms in the PostgreSQL DBMS as part of our overall intrusion detection and response system in a DBMS.

REFERENCES

[1] M .K. Aguilera, R. E. Strom, D. C. Sturman, M. Astley, and T. D. Chandra, "Matching Events in a Content-Based Subscription System, " Proc. Symp. Principles of Distributed Computing (PODC), pp. 53-61, 1999.
[2] Campailla, S. Chaki, E. Clarke, S. Jha, and H. Veith, "EfficientFiltering in Publish-Subscribe Systems Using Binary Decision Diagrams," roc. Int'l Conf. Software Eng. (ICSE), pp. 443-452, 2001.
[3] V. Ganapathy, T. Jaeger, and S. Jha, "Retrofitting Legacy Code for Authorization Policy Enforcement," Proc. IEEE Symp. Security and Privacy, pp. 214-229, 2006.
[4] R. Gennaro, T. Rabin, S. Jarecki, and H. Krawczyk, "Robust andEfficient Sharing of RSA Functions," J. Cryptology, vol. 20, no. 3, pp. 393-400, 2007.

[5] H.-S. Lim, J.-G. Lee, M.-J. Lee, K.-Y. Whang, and I.-Y. Song,"Continuous Query Processing in Data Streams Using Duality of Data and Queries," Proc. ACM SIGMOD, pp. 313-324, 2006.
[6] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, Handbook of Applied Cryptography. CRC Press, 2001.
[7] " Postgresql 8 . 3. The Postgresql Global Development Group " http://www.postgresql.org/, July 2008.
[8] V. Shoup, "Practical Threshold Signatures," Proc. Int'l Conf. Theory and Application of Cryptographic Techniques (EURO1RYPT), pp. 207- 220, 2000.

Advanced Query Evaluation Techniques for Preserving Privacy and Efficiency of Mobile Objects

K.Ruth Ramya[1],K.Priyanka[2],K.Anusha [2],K.Brahmini [2,] G.Mohana Lakshmi[2],K.Abraham[2]

1(Assistant Professor, Dept. of Computer Science And Engineering, KL University.)
2(B.Tech Scholars, Dept. of Computer Science And Engineering, KL University.)

ABSTRACT

This paper presents a novel approach to overcome the difficulty and complexity in addressing the issues of location updating in terms of monitoring accuracy, efficiency, and privacy. Various distressing privacy violations caused by sharing sensitive location information with potentially malicious services have highlighted the importance of location privacy research aiming to protect users' privacy while interacting with Location Based Services. This paper presents a taxonomy of different approaches proposed to enable location privacy in LBS. Location privacy may be obtained at the cost of query performance and query accuracy. The challenge addressed is how to obtain the best possible performance, subjected to given requirements for location privacy and query accuracy. Our proposed framework uses the spacetwist and SCUBA techniques to obtain the privacy and efficiency for continuously moving objects. This approach is flexible, needs no trusted middleware, and requires only well-known incremental NN query processing on the server. This framework offers very good performance and high privacy, at low communication cost thereby providing higher efficiency.

Keywords- **Cloaked region, Clusters, Granular Search, Spatiotemporal.**

I.INTRODUCTION

Every day we witness technological advances in wireless communications and positioning technologies. These developments paved the way to a tremendous amount of research in recent years in the field of real-time streaming and spatio- temporal databases[9,2]. As the number of users of location-based devices (e.g., GPS) continues to soar, new applications dealing with extremely large numbers of moving objects begin to emerge. These applications, faced with limited system resources and near real time response obligation call for new real-time spatiotemporal query processing algorithms[4].Such algorithms must efficiently handle extremely large numbers of moving objects and efficiently process large numbers of continuous spatiotemporal queries. The challenge addressed is how to obtain the best possible performance, subjected to given requirements for location privacy and efficiency at low query evaluation cost.

In a location based service(LBS) scenario, users query a server for nearby points of interest but they may not want to disclose their locations to the service. The benefits of LBS come at the cost of sharing private identity and location information of users with potentially untrusted entities offering such services. Sharing such sensitive information with untrusted servers has recently resulted in various distressing violations of users' privacy. To protect against various privacy threats while using LBS, several studies have proposed different approaches to protect the privacy Intuitively, location privacy may be obtained at the cost of query performance and query accuracy.

A taxonomy of approaches have been used for the location privacy problem.These approaches are based on anonymity/cloaking, transformation and private information retrieval (PIR) techniques. Our paper uses SpaceTwist to   rectifiy the shortcomings of above techniques for k nearest neighbor (kNN) queries. This approach is flexible, needs no trusted middleware, requires only well-known incremental NN query processing on the server and also offers very good performance and high privacy, at low communication cost.

As for efficiency, two dominant costs are: the wireless communication cost for location updates and the query evaluation cost at the database server, both of which depend on the frequency of location updates. Present range and knn queries process and materialize every location update individually there by increasing query evaluation cost. With an extremely large number of objects and queries, this may simply become impossible. In order to reduce the cost i.e to increase the efficiency, here we now propose a two-pronged strategy towards combating this scalability problem. Our solution is based on the fact that in many applications objects naturally move in clusters. We take the concept of moving micro-clusters and exploit this concept towards the optimization of the execution of the spatio-temporal queries on moving objects.We propose the Scalable Cluster-Based Algorithm (SCUBA) for evaluating continuous spatio-temporal queries on moving objects.

II. RELATED WORK

We review existing location privacy protection techniques, which use either spatial cloaking or transformation-based matching.

2.1.Related Work on Preseving Efficiency
2.1.1.Spatial Cloaking

With cloaking, the user location q is enlarged into a cloaked region Q^1 that is then used for querying the server [16]. This way, q is hidden in Q^1. The existing cloaking solutions differ with respect to (i) the representation of Q^1, (ii) the architecture for cloaking, and (iii) the query processing.

Cloaked Region Representation: Cloaked regions come in two forms: they are either plain, connected regions (e.g., rectangles) or they are discrete and posses "multiple parts" (e.g., sets of point locations). Q^1 by a K-anonymous [13] rectangle, which contains the query location q and at least K − 1 other user locations. Figure 1a illustrates a 4-anonymous region Q^1, where u1,u2, and u3 are (4 − 1) user locations.Other work uses circular cloaked regions. The study of Ardagnae al. [17] takes location positioning inaccuracy into account, models the user location as a circular region, and develops several geometric operators for deriving cloaked regions. The cloaked region has also been represented by a point set containing q and a number of dummy locations (generated by the client) . In Figure 1b, q1, q2, q3 are dummy locations, and the cloaked region is Q0 = {q, q1, q2, q3}.

Cloaking Architecture:A simple approach to construct a cloaked region is to do so at the client [16], [17], [11].However, client-based cloaking does not support the use of K-anonymous regions. This requires knowing the locations of other users, which may be achieved by introducing a trusted, third party, location anonymize that knows the locations of a population of users.

(a) K-anonymous   (b) dummies
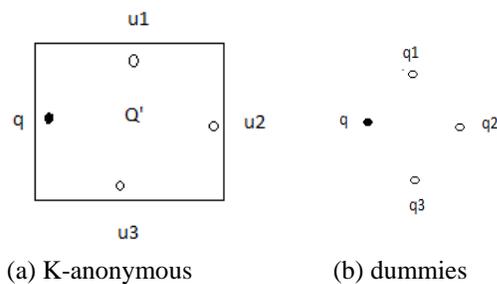
Fig1. Example Cloaked Regions.

Server Side Query Processing: A discrete cloaked region query may be processed by processing each point in the query in turn, returning the union of results[12].In this setting, a negotiation protocol for trading between privacy and result accuracy has also been proposed.The processing of plain cloaked regions is more complex. Specialized

(server-side) algorithms have been proposed for identifying a candidate set that includes the NN for any location in a cloaked region These algorithms go beyond well-known point NN algorithms and introduce complexity.

2.1.2.Transformation-Based Matching

Recently, transformation-based matching techniques have been proposed to enable location privacy. However, these do not offer query accuracy guarantees. A theoretical study on a client-server protocol for deriving the nearest neighbor of q has recently been reported. Its communication cost is asymptotic to pN, where N is the number of POIs. No experimental evaluation of the communication cost and result accuracy of the protocol with real data is available.

2.2.Related Work on Preseving Efficiency

Spatio-Temporal Query Processing: Efficient evaluation of spatio-temporal queries on moving objects has been an active area of research for quite some time. Several optimization techniques have been developed. These include Query Indexing and Velocity Constrained Indexing (VCI) [5],, shared execution[14], incremental evaluation [14], and query- aware moving objects involving high cost.To reduce wireless communication and query reevaluation costs, Hu et. al utilize the notion of safe region, making the moving objects query aware. Query reevaluation in this framework is triggered by location updates only. In this case, the authors combine it with different join policies to filter out the objects and queries that are guaranteed not to join. The limitations of this approach is that the devices may not have enough battery power and memory capacity to perform the complex computations. Our study falls into this category and distinguishes itself from these previous works by focusing on utilizing moving clusters abstracting similar moving entities to optimize the execution and minimize the individual processing[1, 25, 41]. We apply clustering as means to achieve scalable processing of continuous queries on moving objects.

III.FRAMEWORK OVERVIEW

As shown in Fig.2, the PAM[19] framework consists of components located at both the database server and the moving objects. At the database server side, we have the moving object index, the query index, the query processor, and the location manager. The object index is the server-side view on all objects. More specifically, to evaluate queries, the server must store the spatial range, in the form of a bounding box, within which each object can possibly locate. For each registered query, the query index stores: 1) the query parameters (e.g., the rectangle of a range query, the query point, and the k value of a kNN query); 2) the current query results; and 3) the quarantine area of the query. The quarantine area is used to identify the queries whose results might be affected by an incoming location

update. At moving objects' side, we have location updaters.

Without loss of generality, we make the following assumptions for simplicity:

- The number of objects is some orders of magnitude larger than that of queries. As such, the query index can accommodate all registered queries in main memory, while the object index can only accommodate all moving objects in secondary memory. This assumption has been widely adopted in many existing proposals[7], [10], [11].
- The database server handles location updates sequentially; in other words, updates are queued and handled on a first-come-first-serve basis. This is a reasonable assumption to relieve us from the issues of read/write consistency.
- The moving objects maintain good connection with the database server. Furthermore, the communication cost for any location update is a constant. With the latter assumption, minimizing the cost of location updates is equivalent to minimizing the total number of updates.
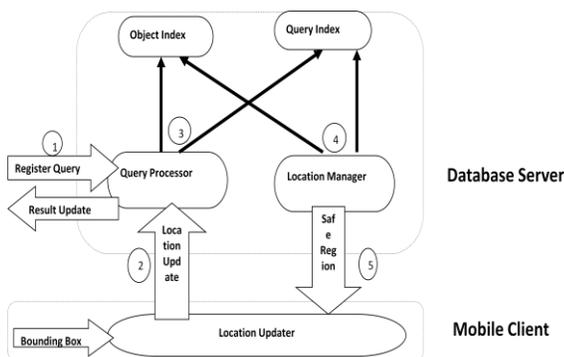
Fig2. PAM framework overview.

PAM framework works as follows (see Fig2): At any time, application servers can register spatial queries to the database server (step _1). When an object sends a location update (step_2), the query processor identifies those queries that are affected by this update using the query index, and then, reevaluates them using the object index (step _3). The updated query results are then reported to the application servers who register these queries. Afterward, the location manager computes the new safe region for the updating object (step _4), also based on the indexes, and then, sends it back as a response to the object (step _5). The procedure forprocessing a new query is similar, except that in step _2 , the new query is evaluated from scratch instead of being reevaluated incrementally, and that the objects whose safe regions are changed due to this new query must be notified.

Algorithm 1 summarizes the procedure at the database server to handle a query registration/ deregistration or a location update.

Algorithm 1: Overview of Database Behavior

1:   while receiving a request do

2:      if the request is to register query q then

3:      evaluate q;

4:      compute its quarantine area and insert it into the query index;

5:      return the results to the application server;

6:      update the changed safe regions of objects;

7:      else if the request is to deregister query q then

8:      remove q from the query index;

9:      else if the request is a location update from object p then

10:    determine the set of affected queries;

11:    for each affected query q^1 do

12:      reevaluate q^1;

13:    update the results to the application server;

14:    recomputed its quarantine area and update the query index;

15:    update the safe region of p;

IV. SPACE TWIST

In this section we present different aspects of privacy related techniques and algorithms. Our proposed framework, called SpaceTwist, rectifies these shortcomings for k nearest neighbor(kNN) queries. Starting with a location different from the user's actual location, nearest neighbors are retrieved incrementally until the query is answered correctly by the mobile terminal.

4.1.The SpaceTwist Client Algorithm

We proceed to present the client-side algorithm for accurate kNN retrieval. We use the notation dist(q, p) to denote the Euclidean distance between two points q and p. The client (i.e., user) executes Algorithm 1 to obtain its k nearest objects from the server (i.e., query processor). The anchor location q^1 is first sent to the server. On the other hand, the user location q is known only by the client. Intuitively, if q and $q^1$are close then few objects are retrieved (i.e., low cost) but less location privacy is achieved. A max-heap W_k, initialized with k virtual objects, maintains the k nearest objects (of q) seen so far. Let γ be the maximum distance in Wk. The demand space is then the circle with radius γ and center q (see Line 3). Let τ be the largest distance to $q^1$of any object examined so

far. The supply space is then the circle with radius τ and center q^1 (see Line 4). Next, the server is requested to return incremental nearest neighbors (INNs) [1] of q^1.

Algorithm 2 Space Twist Client (for kNN query)
algorithm SpaceTwistClient(Value k, Point q, Point q^1)system parameter: packet capacity β

1:    W_k ⟵ new max-heap of pairs <p, dist(q, p)>;

2:    insert k pairs of <NULL,1> into W_k;

3:    γ ⟵ the top distance of Wk; (kth best distance from q)

4:    τ ⟵ 0; (furthest distance seen from q^1)

5:    send an INN query with q^1 to the server;

6:    while γ + dist(q, q1) > τ do

7:    S⟵ get the next packet of points from the server;

8:    τ ⟵ $\max_{p \in S}$ dist(q^1, p); (update supply space)

9:    for all p$\in$ S do

10:   if dist(q, p) < γ then (check demand space)

11:   update W_k (and γ) by using p;

12:   terminate the INN query at the server;

13:   return W_k;

The following discuss a search technique that supports user specified granularities.

4.2. Granular Search
We develop a server-based granular search technique that is capable of retrieving data points from the server with a user specified granularity. This technique enables communication cost reduction and location privacy improvement while providing strict guarantees on the accuracies of the query results. Section 2.1 describes granular search for the case k=1; its implementation is covered in Section 2.2.

4.2.1.Basic Granular NN Search
Recall that the client-side algorithm requests POIs from the server in ascending order of their distance to anchor q1. For the example in Figure 3a, the server returns points in the order: p1, p2, p3, p4. Although p4 is the actual NN of q, it cannot be obtained early by the client.
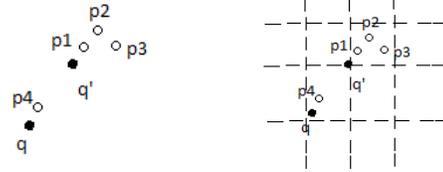
(a) Set of points   (b) Grid cells Fig. 3. Granular Search

The communication cost can be reduced by returning only a sample of the reported POIs. A threshold ϵ is then introduced for controlling the result accuracy.

4.2.2.Implementation of Granular Search
We proceed to consider the implementation of the above method. If the error bound ϵ is given in advance, then it is possible to preselect a data point from each (non-empty) cell and index those points by another (small) R-tree, which is then used at query time. This pre-computation approach becomes impractical when different users use different values for ϵ and may choose these values at run time. In the context of data streams, efficient main-memory data structures for maintaining relaxed results for NN queries with fixed error bounds have been proposed . We are unable to use these because (i) we deal with large, disk-based point sets, and (ii) they require the error bound to be known in advance. Algorithm 2 shows our granular incremental NN algorithm, which takes the user-specified error bound ϵ as input. A conceptual grid with cell extent λ $(=\epsilon/\sqrt{2})$ is imposed on the returned points during runtime. The algorithm also takes an R-tree R (of the data points) and an anchor q^1 as arguments. The notation mindist(q1, e) (maxdist(q^1, e)) represents the minimum (maximum) possible distance between q^1 and an Rtree entry e [1], [1]. Next, $C_\lambda(p)$ denotes the cell containing point p.The algorithm applies INN search [1] around anchor q^1, with two modifications: (i) a set V is employed (Line 3) for tracking the grid cells of the reported points (Line12),and (ii) only qualifying entries that are not covered by the union of cells in V are further processed (Line9).

Algorithm 3 Granular Incremental NN
algorithm GranularINN(R-Tree R, Point q^1, Value ϵ)

1: λ ⟵ $\epsilon/\sqrt{2}$;

2: H ⟵ new min-heap (mindist to q^1 as key);

3: V   new set; (cells of reported points)

4: for all entries e $\in$ R.root do

5: insert <e,mindist(q^1, e)> into H;

6: while H is not empty do

7: deheap <e,mindist(q^1, e)> from H;

8: remove each cell c from V satisfying maxdist(q^1, c) < mindist(q^1, e);

9: if e is not covered by the union of cells in V then

10: if e is a point p then

11: report p to the client;

12: V ⟵ V € {C_λ(p)};

13: else

14: read the child node CN^1 pointed to by e;

15: for all entries e^1 € CN^1 do

16: insert <e^1,mindist(q^1, e^1)>into H;

These client-side processing algorithm and a server-side granular search technique supports user-defined (relaxed) query accuracies. SpaceTwist offers systematic support for managing the tradeoffsamong location privacy, query performance, and query accuracy in mobile services. Empirical studies with real-world datasets demonstrate that SpaceTwist is capable of providing high degrees of location privacy as well as very accurate results at low communication cost.

V.SCALABLE CLUSTER BASED ALGORITHM(SCUBA)

The Scalable Cluster-Based Algorithm (SCUBA) is used for evaluating continuous spatio-temporal queries on moving objects. SCUBA exploits a shared cluster-based execution paradigm, where moving objects and queries are grouped together into moving clusters based on common spatio-temporal attributes. Then execution of queries is abstracted as a join-between clusters and a join-within clusters executed periodically (every time units). In join-between, two clusters are tested for overlap (i.e., if they intersect with each other) as a cheap pre- filtering step. If the clusters are filtered out, the objects and queries belonging to these clusters are guaranteed to not join at an individual level. Thereafter, in join-within, individual objects and queries inside clusters are joined with each other. This two-step filter-and-join process helps reduce the number of unnecessary spatial joins.

5.1.The Notion of Moving Clusters
A moving cluster abstracts a set of moving objects and moving queries. We group both moving objects and moving queries into moving clusters based on common spatiotemporal properties i.e., with the intuition that the grouped entities travel closely together in time and space for some period. We consider the following attributes when grouping moving objects and queries into clusters: (1) speed, (2) direction of the movement (e.g., connection node on the road network), (3) relative spatial distance, and (4) time of when in that location. Moving objects and

queries that don't satisfy conditions of any other existing clusters form their own clusters, single-member moving clusters.
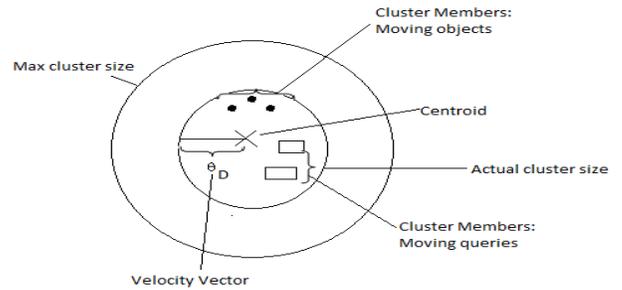
Fig4:Moving cluster in SCUBA

5.2.The SCUBA Algorithm
SCUBA execution has three phases: (1) cluster prejoin maintenance, (2) cluster based joining, and (3) cluster post-join maintenance as depicted in Fig. 5.The cluster pre-join maintenance phase is continuously running where it receives incoming information from moving objects and queries and applies in memory clustering. In this phase, depending on the incoming location updates, new clusters may be formed, "empty" clusters may be dissolved, and existing clusters may be expanded. The clusterbased joining phase is activated every time units where join-between and join within moving clusters is executed. The cluster post join phase is started by the end of the joining phase to perform a cluster maintenance for the next query evaluationtime.
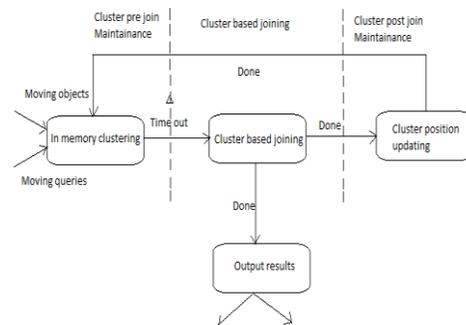
Fig5. State Diagram

Algorithm4 SCUBA()

1: loop

2: //*** CLUSTER PRE-JOIN MAINTENANCE PHASE ***

3: T_{start} = current time //initialize the execution interval start time

4: while (current time - Tstart) < △ do

5: if new location update arrived then

6:△ Cluster moving object o/query q //procedure described in Section 3.2// expires. Begin evaluation of queries

7: //*** CLUSTER-BASED JOINING PHASE ***

8: for c = 0 to MAX GRID CELL do

9: for every moving cluster m_L € G_c do

10: for every moving cluster m_R € G_c do

11: //if the same cluster, do only join-within

12: if (m_L == m_R) then

13: //do within-join only if the cluster contains members of different types

14: if (($m_L.OID_s > 0$) && ($m_L.QID_s > 0$)) then

15: Call DoWithinClusterJoin(m_L,m_L)

16: else

17: //do between-join only if 2 clusters contain members of different types

18: if (($m_L.OID_s > 0$) && ($m_R.QID_s > 0$)) ||

(($m_L.QID_s > 0$) && ($m_R.OID_s > 0$)) then

19: if DoBetweenClusterJoin(m_L,m_R) == TRUE then

20: Call DoWithinClusterJoin(m_L,m_R)

21: Send new query answers to users

22: //*** CLUSTER POST-JOIN MAINTENANCE PHASE ***

23: Call PostJoinClustersMaintenance() //do some cluster maintenance

Algorithm5   DoBetweenClusterJoin(Cluster   mL, Cluster mR)

1: //Check if two circular clusters m_L and m_R overlap

2:if(($m_L.Loc_t.x-m_R.Loc_t.x$)2+($m_L.Loc_t.y$ $-m_R.Loc_t.y$)2)   <   ($m_L.R - m_R.R$)2 then

3: return TRUE; //the clusters overlap

4: else

5: return FALSE; //the clusters don't overlap

Algorithm4 shows the pseudo code for SCUBA execution. For each execution interval   , SCUBA first initializes the interval start time (Step 3). Before   time   interval expires, SCUBA receives the incoming location updates from moving objects and queries and incrementally updates existing moving clusters or creates new ones (Step 6). When time interval expires (location updating is done), SCUBA starts the query execution (Step 8) by performing join-between clusters and join-within   clusters. If two

clusters are of the same type (all objects, or all queries), they arenot considered for the join-between. Similarly, if all of the members of the cluster are of the same type, no join-within is performed. The join-between checks if the circular regions of the two clusters overlap (Algorithm 2), and join-within performs a spatial join between the objects and queries of the two clusters (Algorithm 3). If join-between does not result in intersection, join-within is skipped.

Algorithm6 DoWithinClusterJoin(Cluster mL, Cluster mR)

1: R = Ø//set of results

2: S_q = Set of queries from m_L U mR //query members from both clusters

3: S_o = Set of objects from m_L U m_R //object members from both clusters

//join moving objects with queries from both clusters

4: for every moving object o_i €So do

5: for every moving query q_j € S_q do

6: spatial join between object o_i with query q_j (o_i q_j)⋈

7: S_r=Set of queries from joining oi with queries in S_q

8: for each Q €S_q do

9: add (Q, o_i) to R

10: return R;

After the joining phase, cluster maintenance is performed (Step 23). Due to space limitations, we don't include the pseudo-code for PostJoinClusters Maintenance().The operations performed during post-join cluster maintenance include dissolving "expiring" clusters and re-locating the "non-expiring" clusters (in the ClusterGrid) based on their velocity vectors for the next execution interval time (i.e.T + .).If at time T +   the cluster △asses its dest△nation node, the cluster gets dissolved.SCUBA combines motion clustering with shared execution for query execution optimization. Given a set of moving objects and queries, SCUBA groups them into moving clusters based on common spatio-temporal attributes. To optimize the join execution, SCUBA performs a two-step join execution process by first pre-filtering a set of moving clusters that could produce potential results in the join-between moving clusters stage and then proceeding with the individual join-within execution on those selected moving clusters. Comprehensive experiments show that the performance of SCUBA is better than traditional grid-based approach where moving entities are processed individually. In particular the experiments demonstrate that SCUBA: (1) facilitates efficient execution of queries on moving objects that have common spatio-temporal attributes, (2) has low

cluster maintenance/overhead cost, and (3) naturally facilitates load shedding using motion clusters while optimizing the processing time with minimal degradation in result quality.

VI.CONCLUSION

This paper proposes a framework for monitoring continuous spatial queries over moving objects. The framework is the first to holistically address the issue of location updating with regard to monitoring accuracy, efficiency, and privacy. We provide detailed algorithms for query evaluation/Re evaluation and for protecting location privacy. Existing location privacy solutions either incur high server load, require specialized server implementations, or produce results without practical guarantees on accuracy bounds of query results. This paper concerns the efficient support for location privacy protection by using space twist algorithm. In this paper, we also proposed a unique algorithm for efficient processing of large numbers of spatio-temporal queries on moving objects termed SCUBA. SCUBA combines motion clustering with shared execution for query execution optimization.

VII.FUTURE WORK

Several promising research directions exist. First, it is relevant to extend the cost model to cover real data distributions, as the current model assumes uniform data and may not accurately reflect the distributions found in real-world data. Second, our proposal considers snapshot k nearest neighbor queries. It is of interest to extend them to support also continuous queries. As future work, we plan to further refine and validate moving cluster-driven load shedding, enhance SCUBA to produce results incrementally and explore further through additional experimentation.

REFERENCES

1. G. R. Hjaltason and H. Samet, "Distance Browsing in Spatial Databases," TODS, 24(2): 265–318, 1999.

2. S. E. Hambrusch, C.-M. Liu, W. G. Aref, and S. Prabhakar. Query processing in broadcasted spatial index trees. In SSTD, pages 502–521, 2001.

3. A. Okabe, B. Boots, K. Sugihara, and S. Chiu, Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd ed. Wiley, 2000. 51(10), 2002.

4. M. F. Mokbel and et. al. Towards scalable location-aware services: requirements and research issues. In GIS, pages 110–117, 2003.

5. S. Prabhakar and et.al. Query indexing and velocity constrained indexing: Scalable techniques for continuous queries on moving objects. IEEE Trans. Computers.

6. Y. Li, J. Han, and J. Yang. Clustering moving objects. In KDD, pages 617–622, 2004.

7. D.V. Kalashnikov, S. Prabhakar, and S.E. Hambrusch, "Main Memory Evaluation of Monitoring Queries over Moving Objects".

8. M. F. Mokbel, X. Xiong, and W. G. Aref. Sina: Scalable incremental processing of continuous queries in spatio-temporal databases. In SIGMOD, pages 623–634, 2004.

9. B. Gedik and L. Liu. Mobieyes: Distributed processing of continuously moving queries on moving objects in a mobile system. In EDBT, pages 67–87, 2004.

10. H. Hu, J. Xu, and D.L. Lee, "A Generic Framework for Monitoring Continuous Spatial Queries over Moving Objects," Proc. ACM SIGMOD, pp. 479-490, 2005.

11. X. Yu, K.Q. Pu, and N. Koudas, "Monitoring k-Nearest Neighbor Queries over Moving Objects," Proc. IEEE Int'l Conf. Data Eng. (ICDE), 2005.

12. M. Duckham and L. Kulik, "A Formal Model of Obfuscation and Negotiation for Location Privacy," in PERVASIVE, pp. 152–170, 2005.

13. H. Kido, Y. Yanagisawa, and T. Satoh, "An Anonymous Communication Technique using Dummies for Location-based Services," in IEEE

14. X. Xiong, M. F. Mokbel, and W. G. Aref. Sea-cnn: Scalable processing of continuous k-nearest neighbor queries in spatio-temporal databases. In ICDE, pages 643–654, 2005.

15. C.-Y. Chow, M. F. Mokbel, and X. Liu, "A Peer-to-Peer Spatial Cloaking Algorithm for Anonymous Location-based Services," in GIS, pp. 171–178, 2006.

16. M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The New Casper: Query Processing for Location Services without Compromising Privacy," in VLDB, pp. 763–774, 2006.

17. C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati, "Location Privacy Protection Through Obfuscation-Based Techniques,".

18. G. Ghinita, P. Kalnis, and S. Skiadopoulos, "PRIV´ E: AnonymousLocation-Based Queries in Distributed Mobile Systems," in WWW, pp. 371–380, 2007.

19. Haibo Hu, Jianliang Xu, Senior Member, IEEE, and Dik Lun Lee "PAM: An Efficient and Privacy-Aware Monitoring Framework for Continuously Moving Objects" IEEE ,TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, March 2010.

SINGLE OBJECTIVE RISK- BASED TRANSMISSION EXPANSION

V.Sumadeepthi[1], K.Sarada[2]

[1](Student, Department of Electrical and Electronics Engineering, KL University, India)
[2](Associate Professor, Department of Electrical and Electronics Engineering, KL University, India)

ABSTRACT

This paper proposes a systematic risk based transmission line expansion approach. This expansion method approach comprises of three stages. They are load-driven expansion, security enhancement expansion and risk-based expansion. The main objective of this approach is to minimize the investment cost of the newly added transmission lines by satisfying all the system constraints. In the first stage of transmission expansion the inadequacy problem is identified and it is to be corrected. In the second stage of transmission expansion it is necessary to correct insecurity problem. In the third stage of expansion all the post contingency line overload conditions are eliminated in order to reduce congestion. These three problems are addressed by using Benders decomposition algorithm. The proposed method is illustrated on four-bus and six-bus test system by using MATLAB Software.

Keywords: Transmission Expansion, Adequacy, Security, Risk, Benders- Decomposition, Decision making, N-1 criteria.

1. INTRODUCTION

There is significant evidence that transmission investment in many countries has lagged behind that for load growth and generating capacity additions for some time now [1]. As a result today an increased interest in transmission expansion methods has occurred under deregulation. Adequate transmission capacity is needed to provide security and reliability of the system which are the fundamental needs of modern society. Congestion on a transmission system causes losses and needed extra costs in order to relieve the system from them. Transmission can be a key ingredient in helping to reduce energy prices.

The expansion of the transmission has three purposes: adequacy: to meet future load under normal conditions; security: to meet future load under contingency conditions; risk: to reduce (or eliminate) the need to operate lines at their thermal limits. The third purpose results in relieving congestion in electricity markets that operate based on locational marginal price.

We desire in this paper to develop a transmission expansion method to address all the three purposes. We refer these three as adequacy, security and risk respectively. We know of no systematic approach reported in the open literature addressing these three objectives: adequacy, security and risk. This paper reports on such an approach, which we refer to as risk-based transmission line expansion, which is modular: one may address all three purposes in a systematic fashion or any one of them, or any combination of two of them and here we have eliminate line overloads. An optimization problem is posed to achieve each objective. Each of the three problems utilizes a form Benders decomposition for solution. Benders has been applied to solve power system planning problems before [2][3][4][6] we provide new applications in addressing congestion relief and risk reduction.

2. EXPANSION METHOD

The basic aspects of the expansion method are adequacy, security and risk.

2.1 Adequacy :

The ability of the bulk electric system to supply the aggregate electrical demand and energy requirements of customers at all times, taking into account scheduled and reasonably expected unscheduled outages of system components.

2.2 Security:

The ability of the bulk electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system components or switching operations.

In plain language, adequacy implies that sufficient transmission resources are available to meet projected needs plus reserves for contingencies. Security implies that the power system will remain intact even after outages or equipment failures. This is defined as the ability of the system to operate steady-state-wise within the specified limits of safety and supply quality.

2.3 N-1 Criteria:

Generally the security represents the so-called N-1 criteria. "N" is the total number of transmission "elements" in the system and "N-1" is the total system with one element out of service. The 'minus one' could be a generation unit, a

transmission line or a transmission transformer. The basic idea is that even if one component is lost, the system should still satisfy the load requirements without operating violation. This criterion is used to check the security.

2.4 Risk index:

Within the electric network, an individual disturbance resulting in a severe consequence may occur for a number of reasons at any time. The disturbance may result in overload, voltage collapse, or transient instability, drawing the prevailing system to an uncontrollable cascading situation leading to widespread power outages. Here we consider overloads only. Severity assessment provides a quantitative evaluation of what would happen to the power system in the specified condition in terms of severity, impact, consequence, or cost [11][5]. When the line expansion study is performed, the post contingency consequence to be considered is how much the line flow is approaching the limiting capacity. The plan should be developed in such way that post-contingency flow margin is maintained. We define the severity function for overload as.

$$Sev = M \times \varepsilon \qquad (1)$$

Where

ε is the vector measuring how closely the line flow approaches the rating

M is the penalty vector for the specific operating violation.
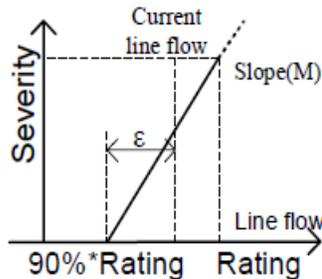
Fig. 1: Overload Severity function

3. FORMULATION AND SOLUTION METHOD

We consider that the three objectives mentioned in the introduction are in fact related problems to be solved together e.g. adequacy expansion has to be done before security expansion):

1. Adequacy expansion: investment decision and adequacy check;

2. Security-enhanced expansion: investment decision and security check;

3. Risk-based expansion: investment decision and risk minimization.

All the above three problems are composed of an economic objective and a reliability sub problem. Four attributes of the overall planning problem drive our choice of solution approach. First, each of the three problems is sequential, where solution to the latter-stage problem (reliability sub problem) depends on determination of the former-stage (economic) objective. For example, in the risk-based expansion, we have to identify the location of the new line before we can evaluate how much risk the system faces. Second, for each of the three problems, the economic objective and the reliability sub problem may be decoupled even though they are related via certain shared variables. Third, the risk evaluation part involves different contingencies, each one effectively a different "scenario" of quantifiable probability. we observe that the investment objectives have integer variables, whereas the reliability sub problems are nonlinear continuous even when using a DC power flow model.

3.1 Benders decomposition

One of the commonly used decomposition techniques in power systems is Benders decomposition. J. F. Benders introduced the Benders decomposition algorithm for solving large-scale mixed-integer programming (MIP) problems. Benders decomposition has been successfully applied to take advantage of underlying problem structures for various optimization problems, such as restructured power systems, operation and planning, electronic packaging and network design, transportation, logistics, manufacturing, military applications, and warfare strategies.

We provide a brief description of the Benders decomposition method, similar to what we have provided in previous publications [11], so that our application to transmission planning is sufficiently self-contained. J.F. Benders introduced the Benders decomposition algorithm for solving large-scale, mixed-integer programming (MIP) problems, which partition the problem into a programming problem (which may be linear or non-linear, and continuous or integer) and a linear programming problem. Problems for which Benders decomposition methods work best are those that have the following structure:

$$Min: z = \underline{c}(\underline{x}) + \underline{d}(\underline{y}) \qquad (2)$$
$$s.t. \ A(\underline{x}) \qquad \geq \underline{b} \qquad (2\text{-}1)$$
$$\underline{E}.\ \underline{x} + \underline{F}\ (\underline{y}) \quad \geq \underline{h} \qquad (2\text{-}2)$$

This problem can be represented as a two-stage decision problem [8]:

Stage 1: Decide on a feasible $\underline{x}$* only considering
$$Min: z = \underline{c}(\underline{x}) + \alpha'\ (\underline{x}) \qquad (3)$$
$$s.t. \ A\ (\underline{x}) \qquad \geq \ \underline{b} \qquad (3\text{-}1)$$

where $\alpha'\ (x)$ is a guess of stage 2 regarding stage 1 decision variable x, which will be updated by stage 2.

Stage 2: Decide on a feasible y^* considering (2-2) given x^* from stage 1.

$$\alpha\,'(\underline{x}^*)=\text{Min: } \underline{d(y)} \qquad\qquad (4)$$
$$\text{s.t.} \quad \underline{F}\,(\underline{y}) \;\geq \underline{h} - \underline{E}.\,\underline{x}^* \qquad\qquad (4\text{-}1)$$

Stage 1 is called the master problem, and stage 2 is called the sub problem. Interaction between stages 1 and 2 and how the problem is solved are shown in Fig. 2. The partition theorem for mixed-variables programming problems [7] provides an important optimality rule on which Benders decomposition is based. If we obtain optimal solution (z^*, x^*) in the first stage and then obtain optimal solution y^* in the second stage, if the upper bound $c(x^*)+d(y^*)$ is equal to the lower bound z^*, then (z^*, x^*, y^*) is the optimal solution for the entire problem. In words, the problem is optimal only when its sub problems are optimal and (2) is satisfied.
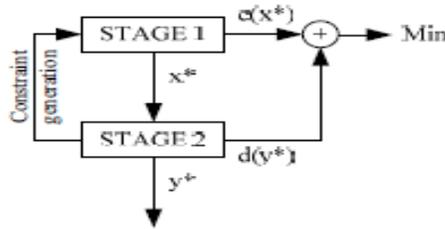
Fig. 2: Solving problem using Benders decomposition

The procedure of Benders decomposition is a learning process (try-fail-try-inaccurate-try-…-solved), and we explain this process as follows. In the left part of Fig. 3, when the stage 1 problem is solved, the optimal value is then sent to stage 2. Stage 2 problem has two steps: 1) Feasibility check sub problem. Check if the optimal solution from stage 1 is feasible. If it is not feasible, the stage 2 problem will send a constraint cut back to stage 1 to let stage 1 remove this infeasible solution set. 2) Optimality check sub problem. Check if the optimal guess of stage 2 from stage 1 is accurate enough. If it is not accurate, a new estimation of $\alpha'(x)$ is sent to stage 1. If the optimal rule is met, the problem is solved.

3.2 DC flow model and decision variable decomposition
As discussed in [9], a linear (DC) power flow relation is often suitable for representing transmission flows in power system planning problems. The decomposed DC flow model with integer decision variable x (1 is build circuit, 0 is not build circuit), which is consistent with formulation [10], is given as follows.

$$\text{f ij} - (\ \gamma\text{ij} -\gamma'\text{ij xij})'(\theta\text{i} - \theta\text{j}) = 0 \qquad\qquad (5)$$

Where
γ Candidate line susceptance
$i\,/\,j$ Bus index

θ Bus angle
f ij   Line flow
If candidate circuit is built xij=1 else we taken as 0.

3.3 Problem formulation: adequacy expansion
For this problem, the target is to minimize the investment under the condition all the load is served.
$$Min \; \underline{C}\,.\,\underline{x} \qquad\qquad (6)$$
$$s.t. \quad UE = 0 \qquad\qquad (6\text{-}1)$$

Where
$\underline{C}$ Cost to build the new line.
$\underline{x}$ Decision variables for the candidate line.
UE   Summation of un served energy.

The sub problem to formulate *UE*, which is feasibility sub problem, also called the minimum load shedding problem (MLS) [13], is as follows.
$$UE= \text{Min } \underline{e}'.\,\underline{r} \qquad\qquad (7)$$
$$\text{s. t.} \quad \underline{s}\ \underline{f} + \underline{g} + \underline{r} = \underline{1} \qquad\qquad (7\text{-}1)$$
$$\text{f ij} - (\ \gamma\text{ij} -\gamma'\text{ij xij})'(\theta\text{i} - \theta\text{j}) = 0 \qquad (7\text{-}2)$$
$$|\underline{f}| \leq (\underline{f}\,max\ \underline{f}'\ max\,.\underline{x}) \qquad\qquad (7\text{-}3)$$
$$|\underline{g}| \leq \underline{g}\ max \qquad\qquad (7\text{-}4)$$

Where
$\underline{r}$ Bus load shedding vector
$\underline{e}$ Bus load shedding penalty vector
$\underline{f}$ Line flow vector
$\underline{s}$ Node-branch incidence matrix
$\underline{g}$ Bus Generation vector
$\underline{1}$ Bus load vector
γij   Existing line susceptance
γij '   Candidate line susceptance
$\underline{f}$ max   Existing line limitation
$\underline{f}$ 'max   Candidate line limitation
$\underline{g}$ max   Generation limitation

3.4 Problem formulation: security-enhanced expansion
After obtaining an adequacy expansion solution, the operating point is feasible under normal conditions. The security-enhanced expansion problem is then performed to ensure that this operating point is feasible under all contingencies. For this problem, the target is to minimize the investment under the condition that all post-contingency line overloads are eliminated.
$$\text{Min } \underline{C}\,.\underline{x} \qquad\qquad (8)$$
$$\text{s.t.} \quad MO = 0 \qquad\qquad (8\text{-}1)$$
Where
$\underline{C}$ Cost to build the new line
$\underline{x}$ Decision variables for the candidate line
MO   *S*ummation of overload under all contingencies
The sub problem to formulate *MO*, which is feasibility check sub problem, called the minimum overload problem, is as follows.
$$MO = \text{Min } \underline{e}'.\ \underline{\eta} \qquad\qquad (9)$$

s.t. $\underline{s}\ \underline{f} = l - g °$ (9-1)

f ij – (γ ij –γ'ij xij)'(θi – θj) = 0   (9-3)

$|\underline{f}| - \underline{n} \le (\underline{f}\,max + \underline{f}'\,max\,.\underline{x})$ (9-4)

Where

$g°$ Generation output (fixed/from load-driven problem)

$\underline{n}$ Overload vector

The minimum overload problem will always feasible for a connected network.

3.5 Problem formulation: risk-based expansion

The objective of the risk-based expansion problem is

$$Min\ C.x + w\ Pk \sum_{k=1}^{N} sev\ k \qquad (10)$$

Where

N   Number of contingencies

w   parameter for weighting risk relative to cost

k   Index of contingencies

Pk   Probability of contingency k

Sevk Severity function of contingency *k*

We do not provide the constraints associated with the complete risk-based expansion problem, but they include power flow constraints, operating limits, and security constraints.

As in the security-enhanced expansion problem, the target of the risk-based expansion problem is to minimize the investment under the condition that all post-contingency line overloads are eliminated, i.e.,

Min $\underline{C}$. $\underline{x}$ (11)

s.t. MO $=0$ (11-1)

where nomenclature is the same as in (8) and (8-1). However, whereas the sub problem of the security-enhanced expansion minimizes composite overload (MO) over all contingencies, here we minimize composite overload risk over all contingencies, which is the optimality sub problem, according to

$$Min\ w\ \sum_{k=1}^{N} Pk.sev\ k \qquad (12)$$

s.t $\underline{s}\ \underline{f} + \underline{g} + \underline{r} = \underline{L}$ (12-1)

   f ij – (γij –γ'ij xij)'(θi – θj) = 0   (12-2)

   $|\underline{f}| - \varepsilon \le 0.9\ ((\underline{f}\,max\ \underline{f}'max\,.\underline{x})$ (12-3)

   Sevk=$\sum \varepsilon$ (12-4)

If the three problems are all to be solved together, then (8), (8-1) are redundant with (11), (11-1), and one of these sets may be eliminated. We include (11), (11-1) so that each of the three problem statements is self-contained. Use of this stage identifies least-cost transmission plans to ensure all post-contingency flows are within their circuit's capacity, use of the latter identifies least cost transmission plans to additionally reduce post contingency loadings near to their

capacities, emphasizing contingencies in proportion to their occurrence probability.

3.6 Solution procedure

A flow chart of the solution procedure is given in Fig. 4 and described as follows:

Step 1: Check if the system is adequate; if yes go to Step 3.

Step 2: Perform the adequacy expansion to find a feasible operation point (6).

Step 3: Check if the system is secure, that is, check if the operating point found in Step 2 is feasible under contingency conditions. If yes go to Step 5.

Step 4: Perform the security-enhanced expansion (8).

Step 5: Perform risk evaluation (12).

Step 6: Check if the cost plus risk is the minimum according to Benders rule. If not, return to   Step 4.

Step 7: Finish and output results.

The three problems, adequacy expansion, security enhanced expansion, and risk-based expansion, can be solved together, but each may provide useful results in a stand-alone solution. For example, an adequate system may not need the adequacy expansion step and a secured system can directly perform the risk evaluation.

4. ILLUSTRATION

4.1 4-Bus test system

We use a small 4-bus test system from [8], which is shown in Fig. 3. The initial system is inadequate, and so there is to be new generation built at bus 4, necessitating an interconnection and possible system expansion. The generation and load data are in below table I and Table II.
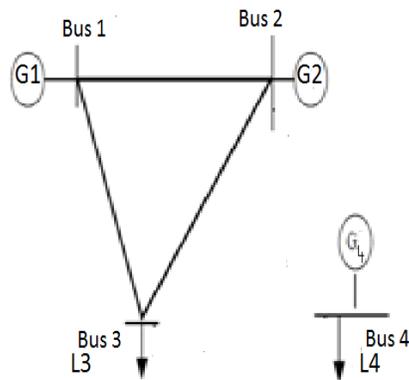
Fig. 3: 4-bus test system

TABLE I
Line Characteristics

Line	From bus	To bus	Susceptance	Capacity	Cost in $
1	1	2	10	200	-
2	1	3	10	200	-
3	2	3	5	200	-
4	2	4	5	150	6,000,000
5	3	4	5	150	5,000,000

TABLE II
Generation Capacity and Loads

Bus	Generation Capacity	Load
1	150	0
2	200	0
3	0	200
4	100	200

4.2 Adequacy expansion

This procedure gives us an adequate system i.e; all the load should be met under normal conditions. The minimum investment needed to make the system adequate is 11,000,000 $. One new line between buses 2 and 4 and one line between buses 3 and 4 are installed. It completed in two iterations. The updated system is shown in Fig 4.
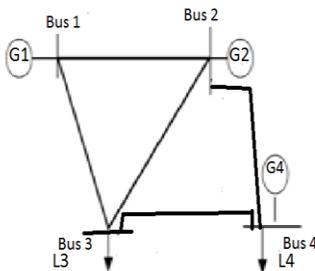
Fig. 4: System after adequacy expansion

4.3 Security enhanced expansion
We need to check if the system is secure or not, i.e., whether all load is met under all considered contingencies,

and if not, we need to identify the minimum cost transmission expansion to make the system secure. The N-1 reliability criterion is used here, so that we consider all contingencies comprised of loss of a single component. It takes 2 iterations to converge. In order that in order to obtain a secure system, an additional investment of 12,000,000 $ is needed. Two additional lines are added between buses 2 and 4 .The updated system is as follows in Fig. 5
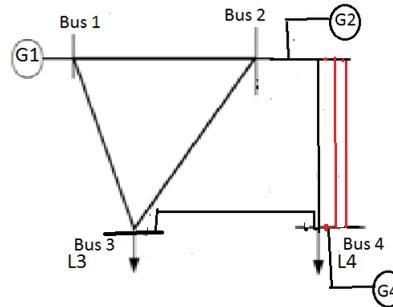
Fig. 5: System after security enhancement expansion

4.4 Risk-based expansion
In this subsection, we use the risk-based expansion to determine additional investment necessary to reduce congestion, i.e., to reduce high post-contingency loadings, emphasizing effects in proportion to the occurrence probability of the contingency which caused them. Contingency probabilities are given in Table III.

TABLE 3
Contingency probabilities

Line	Probability
(1,3)	0.013
(2,3)	0.005
(2,4)	0.015
(3,4)	0.011

It takes 3 iterations to converge and weight w is equal to 6000 with a cost of about $11,000,001.56 in order to minimize congestion of the system. It takes another one line between buses 3 and 4.

The Total Minimum investment necessary to make 4-bus adequate, security and to reduce congestion using this approach is $ 3,4000,001.56.

4.5 6 –bus test system

We use a small 6-bus test system from [5], which is shown in Fig. 5. The initial system is inadequate, and so there is to be new generation built at bus 6, necessitating an interconnection and possible system expansion. The generation and load information are shown in Table I, and all line information is in Table II, for both existing and candidate circuits.
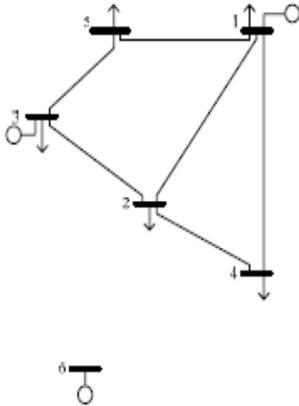
Fig. 6: The 6-bus test system

TABLE 4
Generation Capacity and Loads

Bus	Generation Capacity	Load
1	150	80
2	0	240
3	360	40
4	0	160
5	0	240
6	600	0

TABLE 5

Line   characteristics

Line	Cost(pu$)	Susceptance	Capacity
(1,2)	40	2.50	55
(1,3)	38	2.63	100
(1,4)	60	1.67	30
(1,5)	20	5.00	65
(1,6)	68	1.47	70
(2,3)	20	5.00	110
(2,4)	40	2.50	75
(2,5)	31	3.22	100
(2,6)	30	3.33	100
(3,4)	59	1.69	82
(3,5)	20	5.00	95
(3,6)	48	2.08	100
(4,5)	63	1.59	75
(4,6)	30	3.33	100
(5,6)	61	1.64	78

After adequacy expansion   all the load can be met under normal conditions. The updated system is shown in Fig 7.

Fig. 7: System after adequacy expansion

4.6 Adequacy Expansion

Here our main objective is to minimise the cost under adequacy condition. The minimum investment needed to make the system adequate is pu$130. An additional two duplicated lines between buses 3 and 5 are installed. One new line between buses 2 and 6 and two lines between buses 4 and 6 are installed. The problem takes 4 iterations to converge. This procedure i.e; adequacy expansion gives us an adequate system, i.e., a system for which all load can be met under normal conditions.

4.7 Security enhanced expansion

We need to check if the system is secure or not, i.e., whether all load is met under all considered contingencies, and if not, we need to identify the minimum cost transmission expansion to make the system secure. To obtain a secure system, an additional investment of Pu$50 is needed. Two additional lines are added between buses 2 and 3 and between buses 4 and 6. It takes 6 iterations to converge. The updated system is as follows.
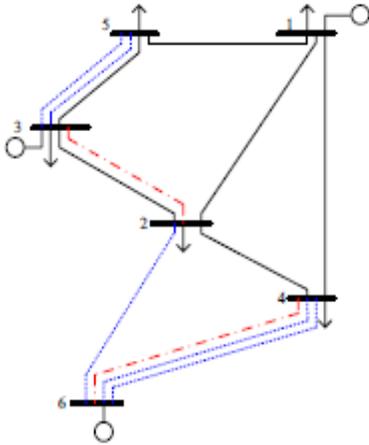
Fig. 8 : System after security enhancement expansion

4.8 Risk-based expansion

In this subsection, we use the risk-based expansion to determine additional investment necessary to reduce congestion. The minimum investment needed to make system to reduce congestion is about   pu$ 30.72 at the 2nd iteration. Contingency probabilities are given in Table 6

TABLE 6
Contingency probabilities

Line	Probability
(1,2)	0.011
(1,4)	0.020
(1,5)	0.014
(2,3)	0.005
(2,4)	0.015
(2,6)	0.022
(3,5)	0.015
(3,6)	0.012

The Total Minimum investment necessary to make 6-bus adequate, security and to reduce congestion is pu$ 210.72.

5. CONCLUSION

This Paper presents a single objective risk-based transmission line expansion approach. The   approach identifies a least cost transmission expansion plan and it can answer how to expand an   inadequate system to adequate one ; how to expand an   insecure system to an secure one; and how to manage post contingency risk. Benders Decomposition is used to solve the problem. The

results are illustrated on a 4-bus and 6- bus test system using MATLAB Programming.

REFERENCES

[1]   Eric Hirst, "U.S. transmission capacity: present status and   future   prospects,"   August,   2004.   [Online]. Available: http://www.electricity.doe.gov/documents/transmission_capacity.pdf
[2]   R. Romero and A. Monticelli, "A hierarchical decomposition approach for transmission network expansion   planning,"   *Power   Systems,   IEEE Transactions on,* vol. 9, no. 1, pp. 373–380, Feb. 1994.
[3]   G. Oliveira, A. Costa, and S. Binato, "Large scale transmission network planning using optimization and heuristic   techniques,"   *Power   Systems,   IEEE Transactions on*, vol. 10, no. 4, pp. 1828–1834, Nov. 1995.
[4]   S. Haffner, A. Monticelli, A. Garcia, J. Mantovani, and R. Romero, "Branch and bound algorithm for transmission system expansion   planning using a transportation model," Generation, Transmission and Distribution, *IEE Proceedings*-, vol. 147, no. 3, pp. 149–156, 2000.
[5]   Risk-based Transmission Expansion   Yuan Li, *Student Member, IEEE, James. D. McCalley, Fellow, IEEE, Sarah Ryan,IEEE 2008*
[6]   S. Binato, M. Pereira, and S. Granville, "A new benders decomposition approach to solve power transmission network design problems," *Power Systems, IEEE Transactions on*, vol. 16, no. 2, pp. 235–240, May 2001.
[7]   J.F. Benders, "Partitioning prodedures for solving mixed-variables programming problems, Numerische Mathematik 4: 238–252, 1962.
[8]   Shahidehopour, M. and Yong Fu, "Benders decomposition: applying Benders decomposition to power systems," *Power and Energy Magazine, IEEE Volume 3*, Issue 2, March-April 2005 Page(s):20 – 2
[9]   Villasana, R.; Garver, L.L.; Salon, S.J., "Transmission Network Planning Using Linear Programming," *IEEE Transactions on Power Apparatus and Systems* Volume PAS-104, Issue 2, Feb. 1985 Page(s):349 – 356
[10] A.M. Geoffrion, "Generalized Benders Decomposition," *Journal of Optimization Theory and Applications,* Vol. 10, no.4, pp. 237-261, 1972.
[11] Ming Ni, McCalley, J.D., Vittal, V. and Tayyib, T., "Online risk-based security assessment," *Power Systems, IEEE Transactions on Volume* 18, Issue 1, Feb. 2003 Page(s):258 – 265.

RANKING AND SUGGESTING POPULAR ITEMSETS IN MOBILE STORES USING MODIFIED APRIORI ALGORITHM

P V Vara Prasad[#1] ,Sayempu Sushmitha[*2], Badduri Divya[*3], Gogineni Riharika[*4], Guntur Vijaya Raghu Ram[*5].

#Asst.Professor,Department of Computer Science and Engineering,
K L University, Green Fields, Vaddeswaram, Guntur District, A.P., INDIA. Pincode : 522502
*Students, B.tech, CSE, K L University

Abstract

We considered the problem of ranking the popularity of items and suggesting popular items based on user feedback. User feedback is obtained by iteratively presenting a set of suggested items, and users selecting items based on their own preferences either the true popularity ranking of items, and suggest true popular items. We consider apriori algorithm with some modifications overcoming the complexity that has been seen in other randomized algorithms. The most effective feature of this algorithm is that it reduces the number of database scans and complexity.

1. INTRODUCTION

1.1 TERMINOLOGY:

In this section we first want to introduce the different terms that we were going to use in our paper as fallows.

1.1.1   Ranking: Ranking is giving rank scores to the most popular item by taking user feedback. The most frequently occurring item is given the highest rank score.

1.1.2   Selection: We focus on the ranking of items where the only available information is the observed selection of items. In learning of the users preference over items, one may leverage some side information about items, but this is out of the scope of this paper.

1.1.3   Imitate: The user study was conducted in very famous mobile stores and which has been used to set of mobiles. The user may check the list and select the set of mobiles which they like most and depending on those like results the new suggestion list has been developed by the algorithm.

1.1.4   Popular: In practice, one may use prior information about item popularity. For example, in the survey the user may select the suggested mobile or they may also select the others. If they selected the already suggested items they will become more popular and if he don't they may get out of the popular list.

1.1.5   Association Rule: Association Rules are if/then statements that help uncover relationships between seemingly unrelated data in the relational database or other information repository. An example of an association rule would be **if a customer buys a nokia mobile, he is 70% interested in also purchasing nokia accessories**.

2. THEORETICAL STUDY

We consider the mobile phone selection and suggesting the best sold mobile and their combinations that were most liked by most of the users. Consider a set of mobiles M: (m1, m2, m3, m4, ...mn) where n > 1. Now we were calculating the set of items in C where were mostly sold and mostly liked by the users, as S

S: (s1, s2, s3, s4, ... sg) where g > 1.

We need to consider an item I, we interpret si as the portion of users that would select item i if suggestions were not made. We assume that the popularity rank scores s as follows:

a)   Items of set S were estimated to is as $s1 \geq s2 \geq s3 \geq$ .... sc,
b)   s is completely normalized such that it is a probability distribution, i.e., $s1 + s2 + s3 + ... + sc = 1$.
c)   si is always positive for all items i.

3. PROPOSED ALGORITHM AND STUDY

We have some of the systems already existing in the same field and we have also identified some of the disadvantages in them as follows:

- The popularity for any item is given based on the production of that item. This may not give good result because customers may not have the knowledge of true popularity they needed and depend on the results given by the producer.

- The updates are performed regardless of the true popularity by virtual analysis.

- Producer have to analyse things manually and complexity involves in this. Due to this time consumption may be high.

- The algorithms used in this system may fail to achieve true popularity.

We consider the problem learning of the popularity of items that is assumed to be apriori unknown but has to be learned from the observed user's selection of items. We have selected a mobile market and mobile distribution outlets as our data set and examined them completely in all areas where we can give the list of items suggested by the users and we have made web-application to make an survey at real-time and considered the data given by more that 1000 members of different categories of people and applied our proposed apriori algorithm on the set of data and started suggesting the item in the mobile outlets for the actual users, which had helped the mobile phone companies and also the outlet in-charges. We have implemented the same in some of the mobile outlets in INDIA where we got very good results. The actual goal of the system is to efficiently learn the popularity of items and suggest the popular items to users. This was done to the user to suggest them the mostly used mobiles and their accessories, such that they also buy the best and at the same time the outlet owner will also get benefited. The most important feature in our project is suggesting the users by refreshing the latest results every time the user gives the input and changes his like list.

Now we have overcome many of the disadvantages of the existing systems and achieved many advantages with the proposed algorithm and method as follows:

- In our approach, we consider the problem of ranking the popularity of items and suggesting popular items based on user feedback.

- User feedback is obtained by iteratively presenting a set of suggested items, and users selecting items based on their own preferences either from this suggestion set or from the set of all possible items.

- The goal is to quickly learn the true popularity ranking of items and suggest true popular items.

- In this system better algorithms are used. The algorithms use ranking rules and suggestion rules in order to achieve true popularity.

4. PROPOSED ALGORITHM
APRIORI ALGORITHM:

This is to find frequent item-sets using candidate generation. Apriori employs an iterative approach known as level-wise search, where k-itemsets are used to explore (k+1) itemsets. First, the set of frequent 1-itemsets is found by scanning database to accumulate the count for each item and collecting those items that satisfy minimum support. The resulting set is denoted by L1. Next, L1 is used to find L2,the set of frequent 2-itemsets,which is used to find L3 and so on, until no more frequent k-itemsets can be found. The finding of each Lk requires one full scan of database.

To improve the efficiency of the level-wise generation of frequent itemsets, an important property is called apriori property, which is used to reduce search space.

Apriori property:

➢ All non empty subsets of a frequent itemset must also be frequent.

➢ Antimonotone property-if a set cannot pass a test, all of its supersets will fail the same test as well.
Apriori algorithm is a two step process, the two steps are
- Join step
- Prune step

Join step:
To find Lk, a set of candidate k-itemsets is generated by joining Lk-1 with itself. This set of candidates is denoted by Ck. Let l1 and l2 be itemsets in Lk-1.The notation li[j] refers to jth item in li. By convention, Apriori assumes that items within a transaction or itemset are sorted in lexicographic order. For the (k-1) itemset, li, this means that the items are sorted such that li[1]<l1[2]<………..<li[k-1].The join Lk-1∞Lk-1, is performed, where members of Lk-1 are joinable if their first (k-2) items are in common.

Prune step:
Ck is a superset of Lk, that is, its members may or may not be frequent but all of frequent k-itemsets are included in Ck. A scan of database to determine the count of each candidate in Ck would result in the determination of Lk. Ck however can be huge. so this could involve heavy computation. To reduce the size of Ck, Apriori property is used as follows:

Any (k-1) itemset that is not frequent cannot be a subset of a frequent k-itemset. Hence, if any (k-1) subset of a candidate k-itemset is not in Lk-1, then the candidate cannot be frequent either and so can be removed from Ck. This subset testing can be quickly done by maintaining a hash tree of all frequent itemsets.

Improving the efficiency of apriori by proposing it with some variations. Several variations are summarized as follows:

➤ Hash-based technique (hashing itemsets into corresponding buckets):A hash-based technique can be used to reduce the size of candidate k-itemsets, C_k, for k > 1.

➤ Transaction reduction (reducing the number of transactions scanned In future iterations):A transaction that doesn't contain any frequent k-itemsets cannot contain any frequent (k+1) itemsets. Therefore, such a transaction can be marked or removed from further consideration because subsequent scans of database for j-itemsets where j > k, will not require it.

➤ Partitioning(Partitioning the data to find candidate itemsets):A partitioning technique can be used that requires just two database scans to mine the frequent itemsets. It consists of two phases. In phase I ,the algorithm sub divides the transaction of D into n overlapping partitions. If the minimum support threshold for transactions in D is min-sup, then the minimum support count for a partition is min-sup x the number of transactions in that partition. For each partition, all frequent itemsets within the partition are found. These are referred to as local frequent itemsets. In phase II, a second scan of D is conducted in which the actual support of each candidate is accessed in order to determine the global frequent itemsets.

➤ Sampling: The basic idea of the sampling is to pick a random sample S of given data D and then search for frequent itemsets in S instead of D.

➤ Dynamic itemset counting: In a dynamic itemset counting technique, new candidate itemsets can be partitioned into blocks by start points. The technique is dynamic in that it estimates the support of all of the itemsets that have been counted so far, adding new candidate itemsets if all of their subsets are estimated to be frequent.

Algorithm:
Apriori:
  Find frequent itemsets using an iterative level-wise approach based on candidate generation.

Input: D, a database of transactions; min-sup, the minimum support count threshold.

Output: L, frequent itemsets in D.

Method:

```
(1)   L1=find_frequent_1_itemsets(D);
(2)       for(k=2;Lk-1≠Ø;k++)
(3)       {
(4)             Ck = apriori_gen(Lk-1);
(5)          for each transaction t Є D
(6)          {
(7)           C1=subset(Ck,t);
(8)           for each candidate c Є Ct
(9)                 c.count++;
(10)  }
(11)  Lk= {c Є Ck | c.count ≥ min_sup}
(12) }
(13) return L=UkLk;
```

Procedure apriori_gen (Lk-1 : frequent (k-1)-itemsets)

```
(1)       for each itemset l1 Є Lk-1
(2)         for each itemset l2 Є Lk-1
(3)  if(l1[1]=l2[1])^(l1[2]=l2[2])^.......^(l1[k-2]=l2[k-
     2])^(l1[k-1]<l2[k-1]) then
(4)  {
(5)           c=l1∞l2;
(6)           if (has_infrequent_subsets(c,Lk-1)
     then
(7)               delete c;
(8)           else add c to Ck;
(9)   }
(10)      return Ck;
```

Procedure has_infrequent_subset (c:candidate k-itemset;

Lk-1:frequent(k-1) –itemsets);

```
(1)       for each (k-1) –subset s of c
(2)            If s Є Lk-1 then
(3)                return TRUE;
(4)  return FALSE;
```

5. RESULTS

The above is the best method of ranking and suggesting the best methods in the scenario of mobile phone outlets in INDIA, which is shown in the following diagram:

Data flow chart of proposed system:

SUGGESTION SET

USER

ENTIRE SET
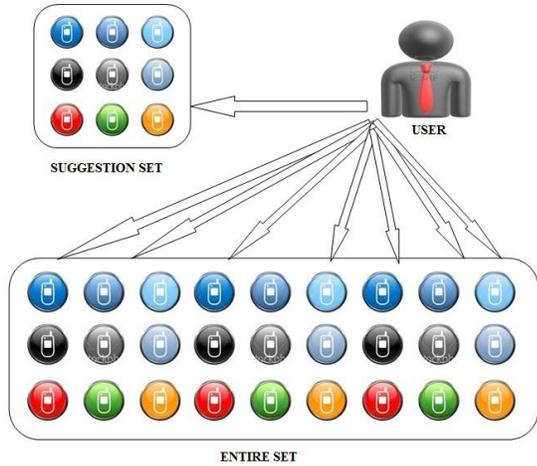
As it was shown in the above diagram we were going to take the most liked items from the users and suggesting the best mobiles or the best set of suggestions that the most of the users liked or ordered.
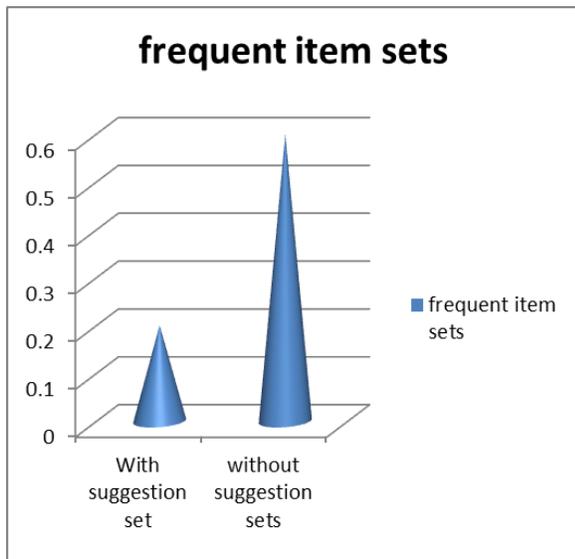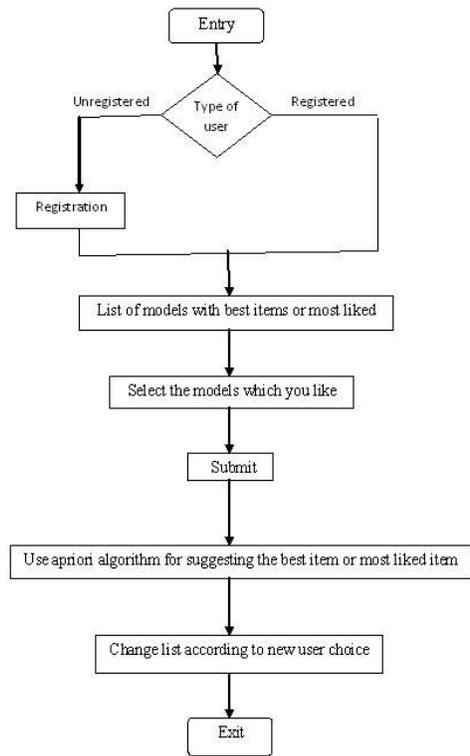
The confidence of the suggestions were also proved by an traditional confidence calculations as follows

In this section we are going to discuss about algorithms. Till now we have discussed some ranking rules , suggestion rules and Frequency move2set algorithm. We have some problems with these, so we go for an algorithm which suits our requirements well. The algorithm is Apriori algorithm. In order to know these algorithms we need to know some concepts of data mining.

Frequent itemsets:

Let I={I1, I2, I3,…., Im} be a set of items. Let D, the task-relevant data, be a set of database transactions where each transaction T is a set of items such that T is a subset of I. Each transaction is associated with an identifier, called TID. Let A be a set of items. A transaction T is said to contain A if and only if A is a subset of T. An association rule is an implication of the form $A \geq B$, where A is subset of I, B is subset of I and A∩B =Ø. The rule $A \geq B$ holds in the transaction set D with support s, where s is the percentage of transactions in D that contain AUB. This is taken to be the probability ,P(AUB).The rule $A \geq B$ has confidence c in the transaction set D, where c is the percentage of transactions in D containing A that also contain B. This is taken to be the conditional probability, P(B/A). That is,

Support(A=>B) = P(AUB)

Confidence(A=>B) = P(B/A)

Rules that satisfy both a minimum support threshold (min_sup) and a minimum confidence threshold (min_conf) are called strong. The occurrence frequency of an itemset is the number of transactions that contain the itemset. This is also known, simply as the frequency, support count,or count of the itemset. The set of frequent k-itemset is commonly denoted by Lk.

$$confidence(A \geq B) = P(A/B) = support(AUB) /$$

$$support(A) = support_{count(AUB)} / support_{count(A)}.$$

Mining frequent itemsets:

In general, association rule mining can be viewed as a two-step process:

1. Finding all frequent itemsets: By definition, each of these itemsets will occur at least as frequently as a predetermined minimum support count, min-sup.

2. Generate strong association rules from the frequent itemsets: By definition, these rules must satisfy minimum support and minimum confidence

6. CONCLUSION

All the previous process already proposed were very complex and contains very complicated computations which made the ranking and suggesting the best and popular items have been more and more complex and not getting to the actual end users. Now we have proposed as very simple randomized algorithm for ranking and suggesting popular items designed to account for popularity bias. This was utilized by many of the mobile outlets in the country successfully.

7. REFERENCES

[1] Huidrom Romesh Chandra Singh, T. kalaikumaran, Dr. S. Karthik, Suggestion of True Popular Items, IJCSE, 2010.

[2] Y.Maanasa, V.Kumar, P.Satish Babu, Framework for suggesting POPULAR ITEMS to users by Analyzing Randomized Algorithms, IJCTA, 2011.

[3] V. Anantharam, P. Varaiya, and J. Walrand, "Asymptotically Efficient Allocation Rules for the Multiarmed Bandit Problem with Multiple Plays—Part i: i.i.d. Rewards," IEEE Trans. Automatic Control, vol. 32, no. 11, pp. 968-976, Nov. 1987.

[4] J.R. Anderson, "The Adaptive Nature of Human Categorization" Psychological Rev., vol. 98, no. 3, pp. 409-429, 1991.

[5] Yanbin Ye, Chia-Chu Chiang, A Parallel Apriori Algorithm for Frequent Itemsets Mining, IEEE, 2006.

[6] Cong-Rui Ji, Zhi-Hong Deng, Mining Frequent Ordered Patterns without Candidate Generation.

[7] Huang Chiung-Fen, Tsai Wen-Chih, Chen An-Pin, Application of new Apriori Algorithm MDNC to Exchange Traded Fund, International Conference on Computational Science and Engineering, 2009.

[8] Milan Vojnovi_c, James Cruise, Dinan Gunawardena, and Peter Marbach, Ranking and Suggesting Popular Items, IEEE, 2009.

UTILIZATION OF TIMESTAMP RECORDS FOR AN HOSPITAL CENTRALIZATION WITH CONSIDERATION OF EMERGENCY CASES

S. Srinivasa Rao[#1], Mandagani Prathyusha[*2], Bathula Surekha Devi[*3], Naraharasetti Jahnavi[*4], Gummadi Sujay[*5],

#Asst.Professor,Department of Computer Science and Engineering,
K L University, Green Fields, Vaddeswaram, Guntur District, A.P., INDIA. Pincode : 522502
*Students, B.tech, CSE, K L University

Abstract

These days, enhancing the quality of healthcare services, such as shortening waiting time and/or providing open-access policy, becomes an important issue even in mid-size hospitals. Even in small to midsize clinics, computerized health care management systems have replaced traditional paper based patient charts and have stored into a database not only patient information but also service-process related information. We introduced some departments like tools, medicine and emergency in order to reduce the patient waiting time deals with the issue of utilizing such per-existing but incomplete data for simulation study. Such an assumption may be justified in cases were data requirement for simulation is precisely defined and all necessary data have been collected according to such requirement. Aiming at the healthcare management systems that maintain the log of operation activities using timestamps. We considered the reduction of patient waiting time as main problem criteria and worked out for it.

Traditional

simulation studies of a hospital centralization is very often assume only till the data centralization and data assumption. Aiming at any of the multi-specialty hospitals of INDIA, to reduce the time wasted for the patients we proposed the present system which can suggest the timestamp that the patient should wait until the doctor will be free for that particular patient. We also designed by consideration of emergency case interruption for the doctors.

1. INTRODUCTION

1.1 TERMINOLOGY:

In this section we first want to introduce the different terms that we were going to use in our paper as fallows.

1.1.1 Time Stamp: A time stamp is a sequence of characters, denoting the date and/or time at which a certain event occurred. A timestamp is the time at which an event is recorded by a computer, not the time of the event is recorded by a timestamp (e.g., entered into a log file)

should be very very close to the time of the occurrence of event recorded.

1.1.2 Open Access Policy: Open access (OA) refers to unrestricted access via the internet to the required hospital website. It is more advanced and very easy method to admit in a hospital now a days even md-size hospitals are providing these services.

1.1.3 Transition Rate: In this particular research paper the world transition rate means the time taken by the patients to get through from one department to other department of the hospital.

1.1.4 Arena: It is a simulation software that is used for the represents of the any/all simulation processes. Arena is a discrete event simulation software simulation and automation software developed by Systems Modeling and acquired by Rockwell Automation in 2000. It uses the SIMA processor and simulation language. As of 2010, it is in version 13.0. It has been suggested that Arena my join other Rockwell software packages under the Factory Talk brand. In Arena, the user builds an experiment model by placing modules (boxes of different shapes) that represent processes or logic. Connector lies are used to join these modules together and timing, the precise representation of each module and entity relative to real-life objects is subject to the modeler.

1.1.5 Service Time: In this particular needed service i.e., to meet the doctor for consultation in the hospitals.

1.1.6 Arrival Rate: The mean number of new calling units i.e., the new patients arriving at a service facility per unit time.

1.1.7 Sensor Networks: Simulation studies of outpatient clinics often involve significant data collection challenges. We describe an approach for data collection using sensor networks which facilitates the collection of a large volume of very detailed patient flow data through healthcare clinics. Such data requires extensive preprocessing before it is ready for analysis. We present a general data preparation framework for sensor network

generated data with particular emphasis on the creation and analysis of patient path strings.

1.1.8   Magnetic Resonance Imaging: MRI is a fairly new technique that has been used since the beginning of the 1980s. The MRI scan used magnetic and radio waves, meaning that there is no exposure to X-rays or any other damaging forms of radiation. An MRI scan is also able to provide clear pictures of parts of the body and spinal cord. Because the MRI scan gives very detailed pictures it is the best technique when it comes to finding tumors (begin or malignant abnormal growths) in the brain. If a tumor is present the scan can also be used to find out if it has spread into nearby brain tissue.

1.1.9   Computed Tomography: A CT scan is a method of taking an image of brain. It is a procedure that produces a clear, two-dimensional image of the brain that shows abnormalities such as brain tumors, blood clots, strokes, or damage due to head injury. A CT scan can help identify the cause of Alzheimer like symptoms either by finding an abnormality or by ruling out certain conditions.

1.1.10   Out Patients: people waiting for consultations or procedures not admitted to hospital are defined as outpatients. Outpatient surgery, also known as ambulatory surgery, same-day surgery or day surgery, is surgery that does not require an overnight hospital stay. The term outpatient arises from the fact that surgery patients may go home and do not need an overnight hospital bed. The purpose of outpatient surgery is to keep hospital costs down, as well as saving the patient time that would otherwise be wasted in the hospital.

1.1.11   Distribution: In mathematical analysis, distributions (or generalized functions) are objects that generalize functions. Distributions make it possible to differentiate functions whose derivatives do not exist in the classical sense. In particular, any locally integral function has a distributional derivative. Distributions are widely used to formulate generalized solutions of partial differential equations. Where a classical solution may not exist or be very difficult to establish, a distribution solution to a differential equation is often much easier. Distributions are also important in physics and engineering where many problems naturally lead to differential equations whose solutions or initial conditions are distributions.

1.1.12   Appointment Scheduling: Based on the patients incoming and outgoing rates and the time available for the doctors the appointment i.e., time given to meet the doctor is scheduled in a hospital. This is called appointment scheduling.

1.1.13   Simulation: Simulation is the imitation of some real thing available, state of affairs, or process. The act of simulating some thing generally entails representing certain key characteristics or behaviors of a selected physical or abstract system. Simulation can be used to show the eventual real effects of alternative conditions and courses of action. Simulation is also used when the real system cannot engaged, because it may not be accessible, or it may be dangerous or unacceptable to engage, or it is being designed but not yet built, or it may simply not exist.

1.1.14   Incoming Logic: There are three parts that we use in the arena software. The incoming logic is a first one and the incoming logic simulates the incoming time rates of outpatients.

1.1.15   Lobby Logic: There are three parts that we use in the arena software. The lobby logic is the second one and the lobby logic simulates the patient movements by interconnecting the incoming logic and the process logic.

1.1.16   Process Logic: There are three parts that we use the arena software. The process logic is the third one and the process logic simulates hospital services information about patient incoming rates, transition rates, and service time are needed for simulation model.

1.1.17   Capacity Planning: Here capacity means the capacity of time or staff and all other modules that effects the patients waiting time in a hospital. Capacity planning is the task of managing the time or staff in order to reduce the waiting time of patients.

1.1.18   Electronic healthcare management system: EHMS is the representation of data in electronic medical records. An electronic medical record (EMR) is a computerized medical record that is created in an organization that delivers care, such as a hospital or physician's office. Electronic medical records tend to be a part of a local stand-alone health information system that allows storage, retrieval and modification of records.

1.1.19   Financial Payoff: The amount necessary to pay a loan in full, with all accrued interest and fees and the prepayment penalty, if applicable. Payoff figures are usually provided to a closing company as correct on a given day. If closing is delayed, the lender has also provided a per diem charge to increase the payoff for every day of delay.

1.1.20   Ultra Sound Scan: Ultra Sound is cyclic sound pressure with a frequency greater than the upper limit of human hearing. Ultrasound is thus not separated from normal (audible) sound based on differences in physical properties, only the fact that humans cannot hear it. Although this limit varies from person to person, it is approximately 20 Kilohertz (20,000 hertz) in healthy, young adults. The production of ultrasound is used in many different fields, typically to penetrate a medium and measure the reflection signature of the medium, a property also used by animals such as bats for hunting. The most

well-known application of ultrasound is its use in sonography to produce pictures of fetuses in the human womb. There are a vast number of other applications as well.

2 THEORETICAL STUDY:

We can see a large number of patients waiting in queues for a long time in hospitals for treatment every day. Service sector has been developing day by day to keep up with the changing world conditions. This development accompanies with planning and management problems. Methods developed for the services provided in hotels, markets, restaurants, factories and hospitals are the new topics of literature. Among these sector is the most reviewed one. There have been rapid changes in the health sector. Several studies have been carried out about hospitals.

Hospital administration striving to provide the best service to the patients with limited staff and equipment imposes some measures to increase the satisfaction and productivity by optimizing the conditions. As technology and science progress, waiting for something causes loss of time for both individuals and institutions. In health sector, patient waiting due to the density causes cost loss.

Patient waiting may also lead progressing of disease and bring social and economic burdens. To minimize this, various measures such as increasing the system working tie or the number of doctors in the system should be taken. Simulation needs data. Collecting data is the key process of simulation. These data cannot be obtained from health units in hospitals. The data used in simulations is non-collectible but available. There are several factors affecting the waiting time in the department. These are insufficient number of junior doctors and working time or greater number of patient admitted to hospital.

3 METHODOLOGY:

To improve resource utilization and to reduce patient waiting time of general hospitals by modifying appointment system, planning the time schedule, and staff assignment. Reduce patient waiting time via appointment scheduling and by open access policy.

We can reduce the patient's waiting time by knowing the service time of each patient as waiting time for a patient is nothing but service times of previous patients. Service time is calculated by using 2 methods, based on two assumptions.

First the service time does not depend on the time or day or the length of waiting queue. Second, a server immediately serves the next patient when its queue is not empty.

3.1 Busy period method:

It is designed for busy periods. We assume that when any patient is waiting in a queue, the server takes no ideal time and immediately serves the first patient in the queue.

$$T^s_{t+1} = T^s_t + S^s_t + I^s_t \text{ -------------- (1)}$$

$$T^s_{t+1} - T^s_t = S^s_t + I^s_t \text{ --------------- (2)}$$

$$T^s_{t+1} - T^s_t = S^s_t \ \ (\text{when } I^s_t = 0) \text{ --- (3)}$$

T^s_t – In time of a patient "t" at server "S"

S^s_t – Service time for patient "t"

I^s_t – Idle time of the server.

Here the idle time of the server is 0 because in busy period method the server mostly will not be in idle time. When identifying the busy period, we use the patient inter-department time (or patient inter-arrival time).

In a busy period, there always are patients waiting for services such that the patient inter departure time will have very little, if any, idle time in it.

3.2 IDLE PERIOD METHOD:

When server operation policy is complex and/or the patient arrival is sparse, the busy period method cannot be used to compute service time distribution. In such cases idle period method is used

We need to trace each movement and calculate the service time by using the timestamps of the patient generated by different servers. In comparison, The busy period method uses multiple patient's time stamps belonging to the target server. When the number of patients being served is small, the availability of servers will be high and patients can move through the series of services without waiting.

In such an idle period, the waiting time can be ignored.

$$T^s_t = D^{s-1}_t + I^s_t \text{ ------------------ (4)}$$

$$T^s_t = D^{s-1}_t \ (\text{when } I^s_t = 0) \text{ ----- (5)}$$

$$S^s_t = D^s_t - T^s_t \text{ ------------------ (6)}$$

$$= T^s_{t+1} - T^s_t$$

$$= D^s_t - D^{s-1}_t$$

T^s_t – In time of a patient "t" at server "S"

S^s_t – Service time for patient "t"

I^s_t – Idle time of the server.

D^s_t – Out time stamp.

4 LIMMITATIONS:

The proposed paper is only considering the mid-size hospitals. All the problems faced by the patients are not solved in this proposed paper. Only some of the

problems are solved such as reduction of patients waiting time and increasing the number of departments.

A short coming of the idle period method is that the number of sample size may result in less accurate estimation of the service time distribution. Thus patients waiting time is calculated and in order to reduce it we need give appointments with respect to certain times that can be allocated for certain patients.
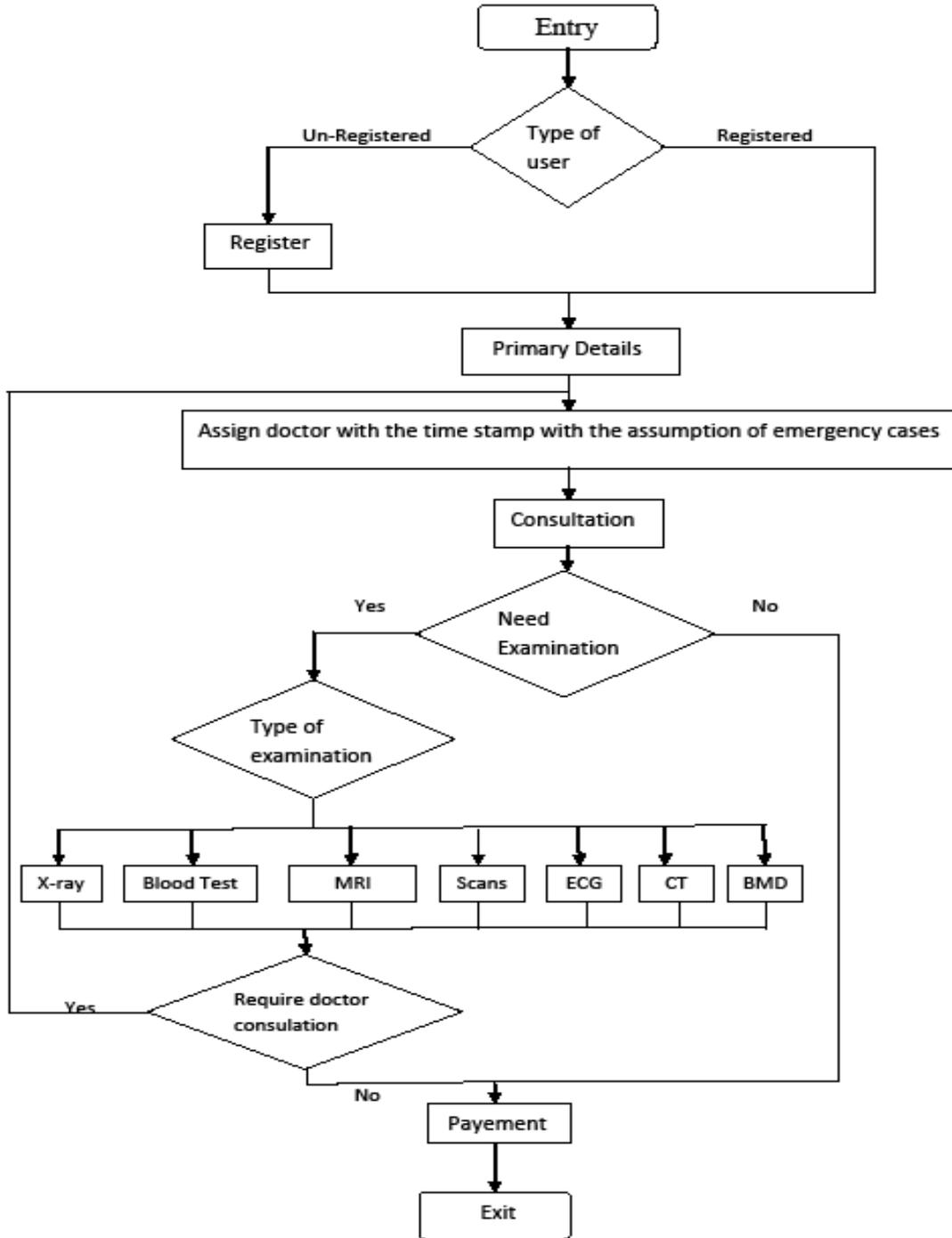
Figure 1: Patient Flow in the Hospital

A patient will enter into the hospital and get register first if he/she is not yet registered, then they will give some primary details about the disease basing on such details we (the hospital staff) have to assign a time slot for the patient. Allotting the slots is the major thing that we need to consider to reduce the patient's waiting time here we have to assume some emergency cases before allocating the slot. It should be make sure that even for the first patient the time should not be 0 (assuming that there will be emergency cases always) and for second (next) patient to go to the same doctor the emergency case time and first patient's time will be added, after that the patient will be send for the consultation if doctor advices for any examinations like X-Ray, Blood test, ECG, Scanning (MRI, CT etc) then patient has to go for the test and if the results demands to go for doctor's consultation again patient has to consult doctor and take correct prescription then go to payment counter after payment patient will exit.

5. RESULTS

5.1 SERVICE TIME DISTRIBUTION:

We need to compute the service time of each process. The service time distribution is derived from the three methods described as busy period method Busy Period Method, Idle Period Method and Emergency Period Method. The estimated service times were presented in Table.

Process	Mean	Variance	Standard Deviation	Methode Use
Check-in	128	20417	142	Busy
Consultation	324	48254	219	Busy
X-Ray	316	7084	84	Busy
Ultrasound	998	86600	294	Idle
CT	2011	43277	208	Idle
MRI	5012	859757	927	Idle
BMD	1517	157010	396	Idle
Payment	200	3056	55	Busy
Lab Test	1215	367348	607	Idle

Table 1. Service Time Distribution (in sec.)

5.2 TRANSITION RATES:

We also need to figure out how each of our patients moves inside the hospital. Using the timestamp created either at the beginning or at the end of each process, we compute the transition rates between processes.

From-To|	Check-in	Consultation	X-Ray	Ultrasound	CT	MRI	BMD	Payment	PT	Lab Test	Exit
Check-in	0	80.60	0.15	0.02	0	0.001	0	32.52	0.79	0.5	2.3
Consultation	3.5	0	52	0.77	0.56	0.85	0.99	87	5.76	3.38	1.7
X-Ray	2.3	67	0	0.13	0.28	0.67	0.12	3.56	45	0.45	0.33
Ultrasound	3.01	16.17	12.41	0	0.13	0.92	1.63	16.98	0.54	23.67	27.44
CT	2.34	65.27	39.21	0.45	0	0.11	5.26	9.09	0.98	1.89	9.32
MRI	12.23	75	6.23	0	0.23	0	1.35	6.23	0.47	6.7	9.5
BMD	0.1	17	79	1.20	0.005	9.45	0	13.55	0.23	4.56	9.5
Payment	0.23	0.76	0.85	0.01	0.97	0.28	0.9	0	48.14	5.51	45.51
PT	7.687	8.53	0.05	0.13	0	0.13	0.005	5.07	0	0.83	84.8
Lab Test	5.9	9.45	10.8	3.56	0.57	5.89	8.65	9.47	8.9	0	57.89
Pharmacy	0.12	69	2.09	0.36	0.5	3.46	2.32	8.64	4	8.1	0.56

Table 2. Transition Rates (in %)

6. CONCLUSION

An effective method of computing the service time distribution from incomplete timestamp information. Proposed method for a case study of a mid-size hospital. The proposed model also consists of adding number of departments for reducing the patients waiting time. The departments added here will be considering the emergency cases of those doctors every time and give the timestamps to the patients in an variant of range of time. So that the wastage of the time of the patient was reduced to maximum extent. In the same way the reduction of some viral and bacterial diseses in the hospital arena was reduced such that the number of patients waiting in the hospital and the time a particular patient staying in the hospital was reduced very much.

7. REFERENCES

[1] Assessing the viability of an open access policy in an outpatient clinic: A discrete event and continuous simulation modeling approach, IEEE, 2009

[2] Yanbing Ju, Aihua Wang, Fengchun Zhu, Analysis of One Hospital Using Simulation, IEEE, 2006.

[3] Vladimir Boginski, In K. Mun, Yuzhou Wu, Katherine P. Mason and Chuck Zhang, Simulation and Analysis of Hospital Operations and Resource, Utilization Using RFID Data, IEEE International Conference on RFID, Gaylord Texan Resort, Grapevine, TX, USA, March 26-28, 2007.

[4] Ren Dawei, Liu Zhaoxi, Zhao Shangwu, Process Analysis of Hospital Outpatient Service Based on Arena, IEEE, 2009.

[5] Semin Sim, Sanju Park, Seogmoon Kim, SeungJae Han, Use of Incomplete Timestamp Records for Hospital Simulation Analysis, IEEE, 2009.

Reduction in Setup Time By SMED A Literature Review

Mr. Rahul.R.Joshi[1], Prof.G.R.Naik[2]

[1]PG. Student, Department of Production Engineering,KIT's College of Engineering,Kolhapur
Shivaji University,Kolhapur(India)
[2] Asso.Professor, Department of Production Engineering, KIT's College of Engineering,Kolhapur
Shivaji University,Kolhapur(India)

ABSTRACT
Manufacturing Organizations faces an a Problem in reduction of cost and efficiency Challenges in their manufacturing Operations. To Stand up in todays Globalization world, Manufacturers need to find ways to reduce Production time and cost in order to improve operating performance and Product quality. This Paper deals with the basic Overview of an a reduction in setup time by SMED. It is normally possible to greatly reduce the setup times and extraordinary results can be Possible through better teamwork,good order,Planning and Simple modifications. The reduction in setup times can be done with help of a SMED Methodology. Each type of Industries can apply the SMED System to reduce their setup times. Single Minute Exchange of Die (SMED) is the approach to reduce Output and Quality losses due to Changeovers.

Keywords:- Changeover, External Setup,Internal Setup, Setup time, SMED

1 INTRODUCTION

In the past a lot of effort has been put to reducing the cycle time and speeding up the output rate whilst totally ignoring the change overtime from one product to another.This has lead to the Economic batch quantity Concept and has resulted in small batches appearing to be Uneconomical to run.[1]

Reducing Setup times (Which we rarely Concentrate on) can give the Equivalent of huge increase in process speed (Which we almost and always concentrate on).This is all achieved without detriment to the quality of the Product. The idea of a setup time reduction Plan is move towards SMED (Single Minute Exchange Die) or OTED (one touch Exchange of dies). [1]

2 DEFINITION OF SETUP TIME/SETUP REDUCTION

Setup time is defined as a the elasped time from when the last part of the current run is Completed until the workcentre starts running the first good piece of the next run[2] or Setup time can be defined as from the stop of Production of Product A until the start of Production of non defective units of Product B.

There are two types of Setup times:- 1)Sequence independent   2)Sequence dependent. [3]

If setup time depends solely on the task to be processed, regardless of its preceding task,it is called a sequence independent. On the other hand, in the sequence dependant type, setup time depends on the both the task and its preceding task. Setup reduction is a Process through which the total time required to changeover or setup equipment or a work centre is dramatically reduced.Through a systematic, Problem solving, Waste eliminating approach to support the movement towards small lot size runs. The main goal of setup reducing is to reduce the downtime of equipment during Changeover.
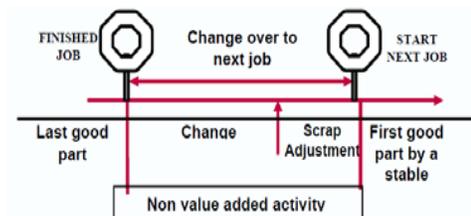
Fig 1: Set Up Reduction representation [4]

As shown in the [5] the Three main reasons for setup reduction are:-

a) Flexibility:- To be able to respond very quickly to changing market demands,you need to be able to produce small lot sizes in an a economical way.

b) Bottleneck Capacities:- Reducing setup times increases the availble capacity,Which can be interesting as an alternatives to buying new equipment or installing an extra shift in situations where the market demand increases.

C) Cost Reduction :- Since, especially on bottlenecks,the direct production cost is related to machine performance, an OEE(Overall Equipment Effectiveness) can easily shown the impact of setup reduction.

3 SMED

Working in any kind of manufacturing environment one of the unfortunate characteristics is waste. Waste can extend from unused raw material to damaged products, and it can carry quite of a financial loss for the company if not treated in an efficient manner. In order to reduce waste, there are several number of methods and strategies that companies can use depending on the desired results. One of the most popular methods is Single Minute Exchange of Die or SMED.

SMED was developed by Shigeo Shingo in 1950s Japan in response to the emerging needs of increasingly smaller production lot sizes required to meet the required flexibility for customer demand. The SMED technique is used as an element of Total Productivity Maintenance (TPM) and "continuous improvement process"[4].It is one of the method of a reducing wastage in a manufacturing Process. The phrase "single minute" does not mean that all changeovers and startups should take only *one* minute, but that they should take less than 10 minutes (in other words, "single-digit minute")

4 BASIC PROCEDURES FOR SETUP TIME REDUCTION

Practical Step 1: Analyze the Setup Operation.

Practical Step 2: Identify the Targets for Improvement.

Sub step 2.1 Eliminate Losses in Setup Operations.

Sub step 2.2 Separate Internal and External Setup.
At this step an important question must be asked for each setup activity. "Do I have to shut the machine down to perform this activity?"The answer helps us in distinguishing between internal and external setup. This step can reduce the setup time by as much as 30 to 50 percent. The three techniques that SMED uses at this step are: Check lists, function checks, and improved transport of dies and other parts.

Sub step 2.3 Convert Internal Setup Steps to External
In order to achieve the single digit setup time objective SMED introduces this step. At this step internal setup activities tried to be converted to external activities. So the total time that the machine is shut down will be reduced. Advance preparation of operating conditions, function standardization, and use of intermediary jigs are the techniques to support the second step.

Sub step 2.4 Shorten Internal Setup and Shorten External Setup Steps.
At this step "specific principles" are applied to shorten the setup times. Implementing parallel operations, using functional clamps, eliminating adjustment and mechanization techniques are used to further setup time reduction

After one round of Sub Steps—from (2.1) to (2.4), return to Practical Steps. Practical Step 3 is next in the procedure sequence

Practical Step 3: Finalize the Improvement Plan.

Practical Step 4: Estimate Postimprovement Setup

Practical Step 5:   Study   and   Evaluate   the Improvement Plans.

Practical Step 6 Arrange the Actual Implementation.

Practical Step 7 Create a Temporary Procedures Manual for the Improved Setup Method.

Practical Step 8 Officially Launch the Improved Setup Method.

Practical Step 9 Implement a "Sideways Expansion" of Single Setup.
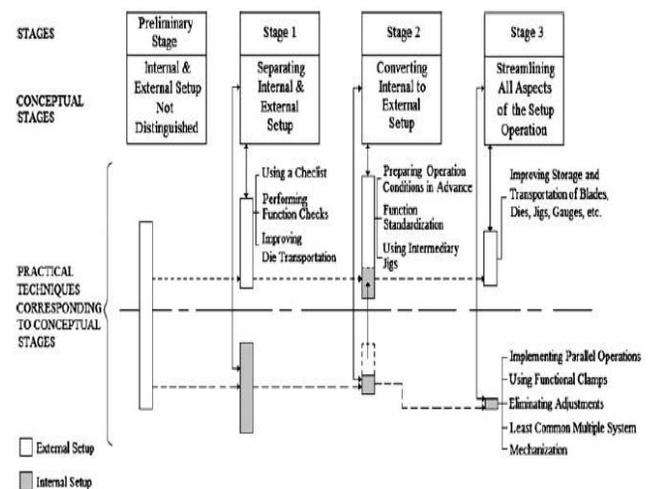
Fig.2 SMED Conceptual Stage and practical Techniques[6]

To make the SMED implementation smoother a group of leveraging tools was also used.They are mentioned in table.[7]

Table I Phases of the SMED conceptand their Leveraging tools

Phases of the SMED concept	Leveraging tools
Phase A: SMED project kick off	1) Analyze the Shop Floor activities in order to differentiate internal from external operations
Phase B: Separate internal from external operations	(2) The use of *checklists* (3) The definition of functions for each worker (4)The improvement of tool transportation
Phase C: Convert internal to external operations	(5) The previous preparation of setup operations (6) The automation of operations (7)The utilization of different tools
Phase D: Improve all aspects of the setup operation	(8) The improvement of tool transportation and warehousing (9) Elimination of settings, calibrations and adjustments (10)The automation of operations

5 BENEFITS [7]

According to Shingo (1985), the main benefits of theSMED application are presented below

Direct Benefits
 - Setup time reduction
- Reduction of time spent with fine tuning
- Fewer errors during change-overs
- Product quality improvement
- Increased safety

Indirect Benefits
- Inventory reduction
- Increase of production flexibility
- Rationalization of tools

6 CONCLUSION

"One of the most noteworthy accomplishments in keeping the price of products low is the gradual shortening of the production cycle. The longer an article is in the process of manufacture and the more it is moved about, the greater is its ultimate cost." (Henry Ford 1926). The changing world economy has caused an increase in the use of just-in-time manufacturing, which results in a trend toward short-run, multiple-product manufacturing.In this Paper the basic ideas of SMED Such as Definition, Basic Procedures for Setup Time Reduction using SMED, SMED implementation smoother a group of leveraging tools and Benefits were Discussed. The importance of setup time reduction was presented using SMED methodologies.After implementing the SMED methodology, it is possible to defend that simple process-based innovations, as the Separation of internal from external operations and the conversion of internal to external operations, are among the key drivers to productivity improvement.

7 References

Books
[1] Setup Time Reduction Process-An Easy Guide(2008)

[2] Maynards Industrial Engineering Handbook.(2004) Tata Mcgrawhill,pp4.81-4.85

Journal Papers
[3] Ali Allahverdi H.M. Soroush
" The Significance of reducing setup times /setup costs" Euporean Journal of Operation Research (2008), pp978-984.

[6] Dr.R.M.Warkhedkar,M.S.Desai,   "Productivity Enhancement by reducing adjustment time and setup change" International Journal of Mechanical Engineering & Industrial Engineering(2011), pp37-42

[7] Antonio Carrizo Moreira,Gil Campos Silva Pais
 "Single Minute Exchange of Die:-A Case Study Implementation Journal of Technology Management and Innoviation(2011)pp29-46

Proceedings Papers
[4] Berna Ulutas
" An Application of SMED Methodology"
World Academy of Science, Engineering and Technology (2011),pp100-103

[5] Dirk V Goubergen,Hendrik v Landeghem
"An Intergated Methodology for More Effective Set Up Reduction"
IIE Solutions 2001 Conference(2001)

Minuscule stride towards ICT

Ms Jinal Jani[1], Ms Manisha Misal[2]

(Asst. Professor Thakur College of Science and Commerce, Mumbai) India

ABSTRACT

ICT (Information & Communication Technology) is based on re-tooling of educational system. The ICT revolution has changed the learning process of childhood up to the real world.

The real challenge is the need to put technology at both , space and ground segment, infrastructure, operations and maintenance system, target group networking and professional management together. Digital learning in India is presently limited, due to which implementation and planning for ICT and development of ICT is important.

Our work suggests that when information and communication technologies (ICT) are adopted by students, how it have impacts on their education. Education becomes more intense, learning process displaced, surveillance increases, ICT have positive effects on learning when taken together with traditional way of education.

Our research shows that students perceive the quality of education is improving as a result of ICT. Based on an analysis of 10 Bombay Municipal Corporation (BMC) school surveys of students in Samatanagar Kandivali, we found that respondents who were most affected by ICT were more satisfied with their studies. Yet, they also perceived that their education had become more effective than those less affected by ICT.

This paper focuses on the influence of school level conditions for the integration and implementation of ICT in education in India.

Keywords – **Analysis**, **Education ,ICT, Methodology, Students.**

I. INTRODUCTION

This paper builds on the BMC Schools in Kandivali. This is the complete analysis of all material collected during visits to 10 schools supported with the depth analysis of data & discussions with students .

1. INFORMATION:

1.1 Training BMC school students of 6th & 7th standard.

1.2 Strength of the students in 6th is 40 and 7th is 40 we provide one student one PC.

1.3 Training timing 1 hour duration for one batch this way two consecutive batches goes on.

2. RELATED WORKS:

2.1 Approaching to the BMC School personally.

2.2 Permission of higher authority with all legal formalities that is letters, list of students name, schedule.

2.3 Taking students from the college gate to the computer lab.

2.4 Allocation of lab special for training students in computer.

2.5 Seating arrangement was done according to the roll number which is assigned by us to students.

2.6 We have also prepared study content in Hindi for effective learning.

2.7 We have maintained attendance muster where we have put up the name list of students, number of practical sessions with dates and timing.

2.8 We have continued this activity till today.

II. METHODOLOGY:

1. The teaching aids which we used is as follows:

1.1 We provided learning text in regional language to the students.

1.2 We have delivered lectures in Hindi with English terminology.

1.3 To make interactive learning session we have use LCD Projector.

1.4 We have also prepared PowerPoint slides for students which content images, pictures, points, explanations, icons which motivate the students towards computer.

1.5 Students perform each practical on computer individually.

1.6   We have designed our course according to the students point of view which has the following topics :

1.6.1   Introduction to Computer.
1.6.2   Uses of Computer.
1.6.3   How to start, access and shut down it?
1.6.4   Paint
1.6.5   Ms-word.
1.6.6   Ms-PowerPoint.

1.7   Students used to write in their book which is one of the resources for them to study .

1.8   We conduct exam after completing the designed course.

1.9   Theory question paper format is TRUE or FALSE, Match the Following, Write in One Sentence which carries 20 marks.

1.10 Practical carries 10 marks which are conducted on Paint, Ms-Word, Ms-PowerPoint.

1.11 Evaluation is done by teachers and suggestions are given on the spot.

1.12 Certificates are distributed in presence of higher authority to motivate students.

III. CHARTS:

Plan to Study ICT in Future				
		2008	2009	2010
6th STD	YES	95	130	150
	NO	90	55	35
7th STD	YES	80	99	140
	NO	82	63	22

Plan to Study ICT in Future

Enrolment of Students from BMC
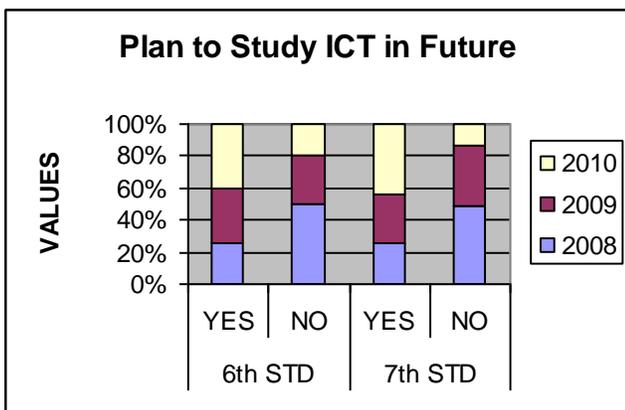
IV. RESULT AND DISCUSSION:

1. Students have benefited with our training course.
2. Students have got border perspective towards latest technology and how to use them in their studies.
3. It is seen that students are more energetic and keen to learn.
4. Students enjoy coming to computer lab. They are waiting for the particular day when they have computer training session.
5. ICT have really made a difference in student's behavior which we can feel and see in them.

V. PROPOSED SOLUTIONS:

1. Instead waiting for the Government funds we came up with this solution to promote ICT in Education in rural areas for rural children's.
2. With the spread of this concept we want to narrow down the gap in digital divide.
3. This will enhance our society towards ICT.

VI. MONITORING AND EVALUATION OF ICTS IN EDUCATION:

This Options Paper articulates key areas for measurement to support the development and delivery of ICT throughout the education sector. The three key areas identified are (1) Infrastructure & Access, (2) Training & Usage, and (3) Impacts. Possible research areas are also presented.

VII. OBJECTIVES:

The aim, of the paper is to contributing to a better understanding of the strengths and weaknesses of using ICT in education. The project deals particularly with identifying methodologies used to favour a use of ICT that in turn fosters added learning value.

VII. CONCLUSION:

More transparent processes for managing school resources. The next most pressing challenge is to increase access in rural areas to Secondary education.
At this level of the education system the
Private sector is growing rapidly and playing the major role of service provider. But in both elementary and secondary education better services will only come about with greater expansion of infrastructure, both within and around schools.

VIII. REFERENCES:

Journal Papers:

1. Neil Anderson, Colin Lankshear, Lyn Courtney, Carolyn Timms.
    An Electronic Journal for leaders in Education.
    http://www.curriculum.edu.au/leader/girls_and_ict_survey:_initial_findings,13812.html?issueID=10270.
    Volume 4 issue 12.

2. The IUP Journal of Information Technology.
    Vol. VII   No.1   March 2011.

3. The IUP Journal of Computer Science.
    The icfai University Press.
    Vol. V   No.4   October 2011.

COLOR AND SHAPE BASED REASSEMBLY OF FRAGMENTED IMAGES

T.Rani Mangammal*, Mrs.G.Prema Priya**

*II M.E CSE, Sri Shakthi Institute of Engineering and Technology, Anna University, Coimbatore
**Assistant professor CSE, Sri Shakthi Institute of Engineering and Technology, Anna University, Coimbatore

Abstract-
The manual execution of reassembling a fragmented image is very difficult. To overcome this difficulty a new technique was introduced in the existing system called "The novel integrated color based image fragments reassembly technique". This technique is divided into four steps, which is based on color. Initially, the spatial adjacent image fragment is discovered. The second operation is to discovery of matching contour segments of adjacent image fragments. The next step is to find the appropriate geometrical transformation for an image fragments contour alignment. Finally the overall image assembly is done via novel based algorithms. In proposing system, the same reassembly technique is applied but additionally the shape alignment algorithm is used to utilize the shape of the fragment contours in order to perform matching. This will produce very satisfactory reassembly results and also it can lead to more efficient, significant reduction in human effort.

Keywords- Fragmented image, spatial adjacent image fragments, color quantization contour segments, geometrical transformation.

1. INTRODUCTION

The problem of reassembling image fragments arises in many scientific fields such as forensics and archaeology. The manual execution of reassembly is very difficult as it requires great amount of time, skill and effort. Thus the automation of such a work is very important and can lead to more efficient, significant reduction in human effort involved. In our work, the automated reassembly of images from fragments follows a four step model, similar to the one presented in [10] for 3-D object reconstruction.

The first step of our approach is, the identification of probable adjacent image fragments, in order to reduce the computational burden of the subsequent steps. There, several color based techniques are employed. This step will produce higher performance. The second step is the identification of the matching contour segments of the image fragments. The corresponding step employs a neural based color quantization approach for the representation of the image contours, followed by a dynamic programming technique that identifies their matching image contour segments. Once the matching contour segments are identified, a third operation takes place. Here, the geometrical transformation, which best aligns two fragment contour along their matching segments, is found. A very popular registration technique is the Iterative Closest Point (ICP) is used to limit the effects of noise. Here in this module we are proposing a new approach of shape alignment method based on the Fourier coefficient.

The last step in solving the fragment reassembly problem is the reassembly of the overall image from its constituent fragments. Here, a novel algorithm is proposed. It is clear that it is essential that each step of the algorithm feeds the next one with correct results; otherwise the image reassembly may contain errors, or may even fail completely. Our goal is to investigate and propose the most robust techniques in order to produce accurate results at each intermediate step. In this paper the integrated method for automatic color based 2-D image fragment reassembly is presented. To summarize, the main steps of the proposed method are shown in Figure. 1.

The main steps are as follows:

1) Discovery of spatial adjacent image fragments
2) Discovery of matching contour segments of adjacent image fragments
3) Image fragments contour alignment
4) Overall image assembly

This paper is organized as follows. Section 2 discusses the related work. Section 3 describes the discovery of the spatial adjacent fragments. Section 4 presents the identification of the matching contour segments of the spatial adjacent image fragments, while Section 5 describes the derivation of the optimal geometrical transformation that aligns contours along their matching segments. Section 6 presents the overall image assembly algorithm. Experimental results are presented in Section 7 and Conclusions are drawn in Section 8.
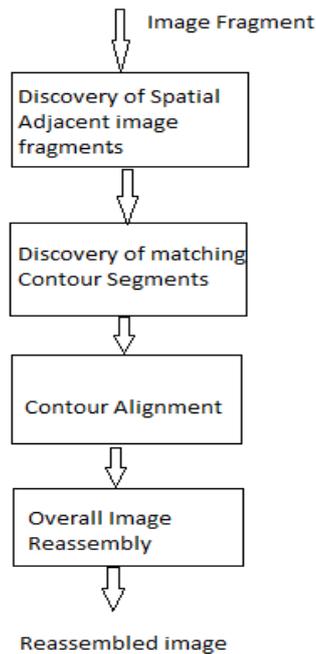
Figure.1. Overall image reassembly approach

2. RELATED WORKS

2.1. Two Dimensional Paper Document Reassembly

Similar to 2-D image fragment reassembly, in paper document reassembly, torn paper fragments must be assembled to form the image of an entire page of a paper document. The above employ shape representations of the paper fragments, in order to reassembly the original documents. In [4], polygonal approximation is initially applied to reduce the complexity of the paper fragment contours and geometrical features are extracted from these polygonal curves. Then, a method based on [7] is used to assemble the entire document from its constituent paper fragments. In [11], shape features, namely turning functions, are estimated from every fracture contour and are utilized to discover matching contour segments. After that, each matching is assigned a confidence score. The alignment transformation of the fragments is simultaneously found during matching. After the discovery of matching contour segments, the final reassembly step performs two actions, namely matching relaxation and fragments merging. During matching relaxation, every pair of aligned input fragments is checked for overlap along their matching contour segments. If they overlap, this matching is discarded. Otherwise, the neighbouring fragments of this pair are identified. A score, called support, is assigned to the neighbourhood of each pair of non-overlapping fragments. This score increases as the number of neighbouring fragments as well as the matching confidence assigned to pairs of fragments increase. The fragments that have neighbourhoods with maximum support are merged into new fragments and the whole procedure starts again, i.e.,

the matching contour segments are identified for all pairs of fragments, and so on.

2.2. Two Dimensional Puzzle Reassembly

Many methods were also proposed for the 2-D puzzle reassembly problem. In [6], color and textural features of the puzzle pieces are utilized. The matching and alignment of puzzle pieces is carried out using an FFT-based image registration technique. In [5], the puzzle reassembly process consists of two steps; frame and interior assembly. A travelling salesman problem (TSP) is formulated for frame assembly, while backtracking and branch&bound techniques are employed for interior assembly. The puzzle pieces are matched employing (L_2) the distance of their contours curves. In [9], the overall puzzle assembly is done using a Best-First procedure. There, two criteria are utilized to sort matching contour segments. The first one is the residual error of corresponding contour pixels after the discovery of the optimal geometrical transformation, while the second criterion is the arc-length of the matching contour segments. It is clear that the problem of 2-D puzzle assembly does not meet the major difficulty of the image and object reconstruction problems; that is the missing or highly damaged image (or) object fragments. Thus, in general, the algorithms proposed in this field would be inadequate to solve such problems.

2.3. Three Dimensional Object Reconstruction

Regarding 3-D object reconstruction, an automatic method for matching and alignment of 3-D, free-form archaeological fragments is proposed in [8]. The input fragments are not pre-processed. The matching is performed utilizing only the 3-D points of the whole surfaces of the objects. The output matching-alignment minimizes the distance between the 3-D surface points of the two fragments. Andrews *et al.* [1] propose an automatic method for the reconstruction of pairs or triplets of 3-D symmetric archaeological fragments. The matching is found through a two-phase method. During the first phase several matching's-alignments are estimated for every pair of fragments. The 3-D points of the fracture curves in the outer and inner surface of the fragments as well as the axis of rotation of the fragments are utilized to this end. In the second phase, these matching's are refined using the quasi-Newton method, and evaluated according to several criteria namely the angle formed by the fragments rotation axes, the perpendicular distance between the rotation axes and the distance of the matched fracture curves points. Eventually, one matching is retained for every pair of fragments. Finally, in the overall object reconstruction step, a greedy merge strategy selects pairs of fragments to form triplets. A human-supervised collaborative reconstruction system is described. The aim of the system is to propose a potential matching between any pair of input fragments. The matching is found by utilizing shape feature estimated from all 3-D points in the fracture curves of the fragments. The shape similarity of the fracture curves is ranked with a cyclic distance algorithm. Each matching defines correspondences between 3-D points in the fracture curves. Then the users

select to merge or not the proposed fragments. The fragments alignment is performed interactively by the users. The object reconstruction procedure follows the merge-update paradigm.

3. DISCOVERY OF SPATIAL ADJACENT IMAGE FRAGMENTS

In this step a basic approach for identifying the spatial adjacent image fragments are presented using content based image retrieval. Color quantization is used to find the normalized quantized color image histograms, which can be used for color image retrieval. The Spatial Chromatic Histogram, which provides information both of color presence and color spatial distribution. The Spatial Chromatic Histogram [3].

$$S_I = (h(i), b(i), \sigma(i)), \; i = \{1, \ldots . C\}$$

In this equation h denotes the normalized color histogram, i.e., **h(i)** is defined as the number of pixels having color i divided by the total number of pixels, **b(i)** is a 2-D vector expressing the center of mass and **σ(i)** is the standard deviation of the color label. Histogram Intersection measures and scaled them to the range [0, 1], with 1 denoting a perfect similarity. The utilized matching measures are the following:

1) Scaled L_1 norm

$$d_{L1}(h_1, h_2) = 1 - 0.5 \sum_{i=1}^{c} |h_1(i) - h_2(i)|$$

2) Scaled L_2 norm

$$d_{L1}(h_1, h_2) = 1 - 1/\sqrt{2} \sum_{i=1}^{c} (h_1(i) - h_2(i))^2$$

3) Scaled Histogram Intersection

$$d_{HI}(h_1, h_2) = \sum_{i=1}^{c} min(h_1(i), h_2(i))(1 - |h_1(i) - h_2(i)|)$$

h_1, h_2 denote the normalized color histograms, extracted from images I_1, I_2 respectively.

4. DISCOVERY OF MATCHING CONTOUR SEGMENTS OF ADJACENT IMAGE FRAGMENTS

In this step a Smith Waterman algorithm is presented in order to match the colors appearing in the contours of adjacent image fragments. Various color similarity criteria are being evaluated. Based on such similarity criteria, for each image fragment, one matching contour segment with other image fragments is retained. In order to avoid comparing directly contour pixel colors that may contain noise, a color quantization pre-processing step is utilized, which takes pixel samples from the contours of all image fragments.

A Kohonen neural network (KNNs) is used for color quantization purposes. KNNs belong in the class of unsupervised neural networks. They can cluster input vectors without any external information, following an iterative procedure based on competitive learning. KNNs consist of two node layers; the input and the output layer. In the former, the number of nodes equals the dimension of input vectors, while in the latter the number of nodes equals the amount of produced clusters. The nodes in the output layer are organized by means of a lattice. In KNNs, each node in the input layer s_i has a connection w_{ik} with every node in the output layer. For a network with n input nodes, the weight vector ending at an output node $w_j = [w_{1k}, w_{2k}, \ldots, w_{nk}]$, is the center of a cluster. In KNNs, given an input vector, the output node with the highest response (winning node) for that as well as all "neighboring" nodes that belong to an area around it, update their weight vectors.

5. IMAGE FRAGMENTS CONTOUR ALIGNMENT

The purpose of this step is to find the appropriate geometrical transformation of one fragment relative to its adjacent one, in order to align them along their matching contour segments. Many variants of the ICP algorithm [2] are employed and evaluated to this end.

The ICP algorithm generally starts with two point sets (contour segments in our case) and an initial guess of their relative rigid body geometrical transformation. It then refines the transformation parameters, by iteratively generating pairs of point correspondences and by minimizing an error metric. Additionally we are proposing a new approach of shape alignment method based on the Fourier coefficient. The aligned fragments of the images will assemble to obtain the required resultant output image.

5.1. Proposing Work

5.1.1. Shape Alignment Using Fourier Coefficient

A new approach of shape alignment method based on the Fourier coefficient will propose to match the contour segments. In this method, the closest contour of a planar object can be represented by a parametric equation γ: [0; 2π] → C

$$l \to x(l) + i\, y(l)$$

With $i^2 = -1$. For latter use, the Fourier Coefficients of γ is given by

$$C_k(\gamma) = \int_0^{2\pi} \gamma(l)\, e^{-ikl}\, dl, \;\; k \epsilon Z$$

Now let γ_1 and γ_2 be centered (according to the center of mass) and normalized arc length parameterizations of two closed planar curves having shapes F_1 and F_2. Hence, scale factor and translation between the two curves can be ignored.

Ghorbel shown that the following quality is a metric between shapes:

$$d(F_1, F_2) = \inf_{(l_0, \theta)} \| \gamma_1(l) - e^{i\theta} \gamma_2(l + l_0) \|$$

Where $T = [0; 2\pi]$ is the range of the rotation angle θ, and of the difference between the starting description points for the two curves l_0. By using the shift theorem in the Fourier domain, computing such a distance comes down to the minimization of

$$f(\theta, l_0) = \sum_{k \in Z} |C_k(\gamma_1) - e^{i(k l_0 + \theta)} C_k(\gamma_2)|^2$$

Persoon & al. proposed a solution to compute l_0 and θ. First, l_0 is one of the zeros of the function

$$g(l) = \sum_k \rho_k \sin(\varphi_k + kl) \sum_k k \rho_k \cos(\varphi_k + kl) - \sum_k k \rho_k \sin(\varphi_k + kl) \sum_k \rho_k \cos(\varphi_k + kl)$$

6. OVERALL IMAGE ASSEMBLY

Once the matching contour segments of couples of input image fragments are identified and properly aligned, the remaining step is the reassembly of the overall image. In this module we are presenting the overall assembling of the image fragments obtained from the previous approach. Since the criteria that are based on the contour matching do not suffice for the overall image reassembly, a novel feature, namely the alignment angles found during the process in module is introduced. Here an integrated method for automatic color based 2-D image fragment reassembly is presented. Consider three image fragments f_i, f_j and f_k each one having one matching contour segment with the rest ones. We denote by Θ_j the rotation angle by which the individual fragment f_i must be rotated, in order to be correctly placed inside the overall reassembled image. The alignment angle, by which we must rotate fragment f_i to align it with the matching contour segment of fragment f_j (before fragment f_j is rotated by Θ_j), is denoted by Θ_{ij}. In order to align fragments f_i and f_j with respect to each other and place them correctly in the reassembled image, the following steps must be performed:

1) Rotate fragment f_j by Θ_j to correctly orient it in the assembled image.

2) Rotate fragment f_i by $\Theta_{ij} + \Theta_j$ to correctly align its matching contour segment with the corresponding matching contour segment of fragment. This procedure will simultaneously align fragment f_i with fragment f_j and provide its correct orientation inside the entire image.

7. IMAGE REASSEMBLY EXPERIMENTS

To evaluate the performance of the reassembly method, the camera image (fig.2.) was taken and it was fragmented into six pieces that were taken as an input for reassembling the image (fig.3.) While executing, the histogram of the original camera image was displayed (fig.4.) and the adjacent images were identified at the very first step.

Figure.2. Camera image

Figure.4. Fragments of the camera image
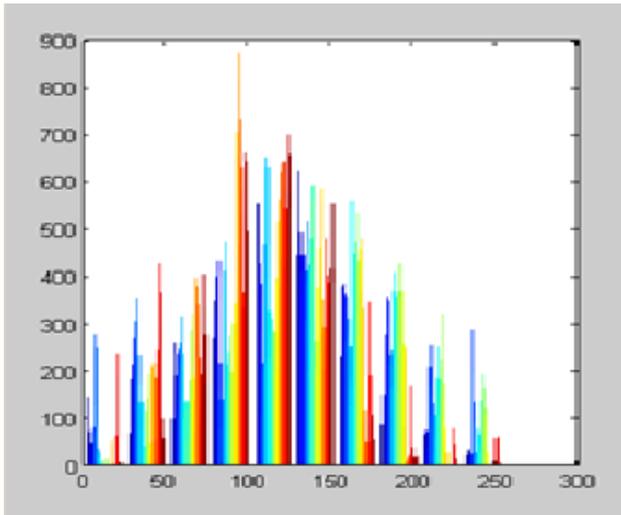
Figure.3. Histogram of the original camera image
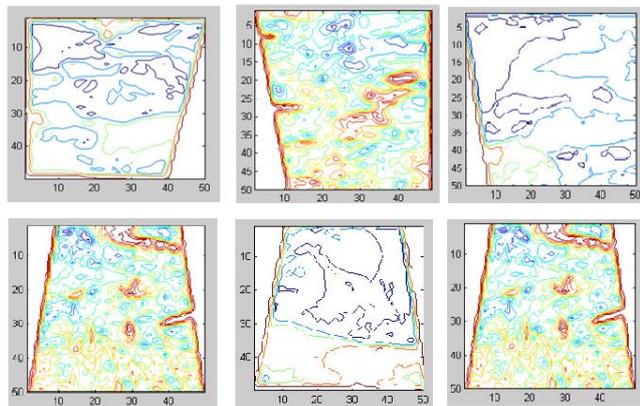
Figure.5. Contour Segments of the fragments

Three sets of the adjacent image were found based on the distance of the color pixels. After that, the matching contour segments were discovered (fig.5. and fig.6.) for each fragment by using smith waterman algorithm. This algorithm was used to match the color appearing in the contours of adjacent image fragments. It must be properly aligned before reassembly of the overall image. For aligning the correct matching contour segments, the Iterative Closest Point algorithm was employed. At the end of the reassembly method, the matched fragments were reassembled (fig.7.)

(1) Fragment 1 and 3

(2) Fragment 2 and 4

(3) Fragment 6 and 5

Figure.6. Alignment of fragments with contour segments

Figure.7. Automatically reassembled image produced from the fragments of fig.4.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a novel integrated color based image fragments reassembly method that consists of several distinct novel algorithms, which produced satisfactory reassembly results. The shape alignment algorithm will be used to improve the matching contour segments of the original camera image fragments.

REFERENCES

[1] S. Andrews and D. H. Laidlaw, "Toward a framework for assembling broken pottery vessels," in *Proc. 18th American Conf. Artificial Intelligence*, 2002, pp. 945–946, American Association for Artificial Intelligence(AAAI).

[2] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image Vis. Comput.*, vol. 10, no. 3, pp. 145–155, 1992.

[3] L. Cinque, G. Ciocca, S. Levialdi, A. Pellicano, and R. Schettini, "Color based image retrieval using spatial chromatic histograms," *Image Vis. Comput.*, vol. 19, pp. 786–979, 2001.

[4] E. Justino, L. S. Oliveira, and C. Freitas, "Reconstructing shredded documents through feature matching," *Fores. Sci. Int.*, vol. 160, pp.140–147, 2006.

[5] Kalvin, E. Schonberg, J. Schwartz, and M. Sharir, "Two dimensional model based boundary matching using footprints," *Int. J. Robot. Res.*,vol. 5, no. 4, pp. 38–55, 1986.

[6] W. Kong and B. B. Kimia, "On solving 2D and 3D puzzles using curve matching," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 2, pp. 583–590.

[7] H. C. G. Leitao and J. Stolfi, "A multiscale method for the reassembly of two dimensional fragmented objects, " *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1239–1251, Sep. 2002.

[8] G. Papaioannou, E. A. Karabassi, and T. Theoharis, "Reconstruction of three-dimensional objects through matching of their parts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 114–124, Jan.2002.

[9] M. S. Sagiroglu and A. Ercil, "A texture based matching approach for automated assembly of puzzles," in *Proc. 18th Int. Conf. Pattern Recognition (ICPR)*, 2006, vol. 3, pp. 1036–1041.

[10] R. Willis and D. B. Cooper, "Computational reconstruction of ancient artifacts," *IEEE Signal Process. Mag.*, pp. 165–183, Jul. 2008.

[11] L. Zhu, Z. Zhou, and D. Hu, "Globally consistent reconstruction of ripped-up documents," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 1, pp. 1–13, Jan. 2008.

STUDY OF EFFECTS OF VIBRATION ON GRIP STRENGTH

[1]Akhilesh.H.Gaidhane ,[2]Dr S. G. Patil

[1] (Reader Mech.engg.deptt, RSR RCET Raipur) [2] (Proff.Mech.Engg Deptt, P.R.M.I.T & R Amravati)

Introduction

Hand Arm Vibration Syndrome (HAVS) - A collective term that includes a number of disease patterns involving the vascular, neurological and musculoskeletal systems. These disease patterns are associated with exposure to hand-arm vibration and are experienced in particular in the hands and forearms of the exposed worker.

Carpal Tunnel Syndrome (CTS) - Carpal tunnel syndrome is a condition caused by compression of the median nerve within the carpal tunnel at the wrist. CTS is characterised by numbness, tingling, burning, or pain in the thumb, index, and middle fingers.Significant exposure to hand-arm vibration – Employees whose exposure to hand-arm vibration represents a risk to their health as determined by a risk assessment. This means:

1) All workers regularly exposed to hand-arm vibration above the action level of 2.5 m/s² .
2) Workers who are only occasionally exposed above the action level but a risk assessment or other factors indicate that the pattern of exposure may pose a risk to health, for example:

a. Use of specific tools or in specific jobs where a detailed risk assessment shows greater than 2.0 m/s².
b. Pragmatic, based on combination of risk assessment, knowledge of tools and uncertainty regarding the frequency or duration of exposure,
c. Any job where the worker experiences numbness or tingling in fingers after 5-10 minutes of continuous use of tool,
d. Jobs where a claims history indicates that HAVS may be a problem.

As Vibration exposure is difficult to assess directly using many fast Fourier (FFT) spectral analyzers because of long task cycle times. Exposure time can-not be accurately estimated using time standards because of high variability between operators and work methods. It is difficult to record vibration without interfering with the operation. Alternately, it is divided into Hand-arm vibration (HAV), affecting workers who use all manner of vibrating pneumatic, electrical. Hydraulic and gasoline powered hand held tools. Due to the weight of the tool and awkward positions that a hand tool operator has to adapt to sometimes, he/she is forced to let the tool rest against his/her torso in an attempt to make the task more comfortable and also to damp the vibration. This results in vibration being transmitted to the body through hand-arm system.

Methodology

For this project a experimental setup is made for identifying and measuring the grip exertion before work while working and after the work of vibration. For this The participants taken were 40 unpaid volunteers (20 men and 20 women) with varied backgrounds in manual work. Participants were not recruited based on their history of work in any particular industry or history of performing specific work tasks. All were in good health at the time of the study, and no participants had acute or chronic musculoskeletal injuries to their upper extremities. Each provided written informed consent before participation. Some of the participants were right handed that is their right had been dominant and some were of left handed that is their left hand is dominant and other non dominant. Different variables are independent and dependent.

Independent variable are

- Vibration level i.e cycles per second and amplitude.
- Dominant and non dominant.
- Male and female participants.
- Peak Grip force in Kgf.
- Average of two values i.e I Maximum voluntary exertion and II Maximum voluntary exertion.

Dependent variables are-

1. effect on grip strength, Kg f
2. Other symptoms

Experimental setup

A vibrator of around 2500-3500cycles per min is developed for this experiment using mechanical attachment attached to a single phase brushed motor of rating 3500rpm make Volco. The vibration is given by connecting rod connected offset to the motor, and is given to the pvc handle attached on the top of the box. This box is connected to a electronic circuit which calculate the voltage , frequency . this experiment is designed for single minute and for more time it can be adjusted as required. For experiment subject is told to hold the vibratory handle and it is started in timer mode. After one minute he machine automatically stops and frequency is displayed on the screen provided. As due to the force exerted by the subjects the frequency changes

for every person so it can't be fixed as per the rating. For this reason the same as been provided in the machine.

Experimental Task

Participants / Subjects

To study the influence of segmental vibration on the workers, 20 participants were taken among them 10 are males and 10 are females . Each participant was in good healthand no participant had acute or chronic musculoskeletal injuries to their upper extremities. during the experiment and had their breakfast during the experiment. participants were not recruited based on their history of work in any particular industry or history of performing specific work task. Grip strength were measured before the experiment, after one minute of vibration then after 3 minutes of vibration and after 5 minutes of vibration. Anthropometric data taken of the subjects is given further.

First the subject is told to take some breakfast and then the experiment is started.

1. Grip strength is recorded using the analog grip strength meter.
2. Subject is told to sit on the chair and hold the vibrator machine handle.
3. The machine is set in timer mode and is started.
4. After one minute it stops automatically and frequency is displayed on the screen and again the grip strength is taken. The same is noted down,
5. subject is told to get ready for next minute reading.
6. This same procedure is repeated for three minute and five minute and the record is noted in the table.

Experimental setup

Electronic circuit box
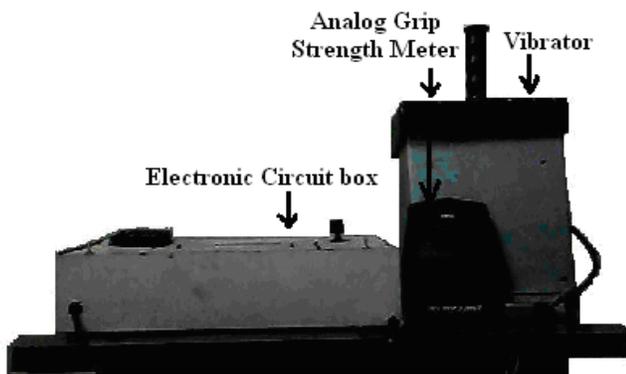
GRIP STRENGTH DYNAMOMETER
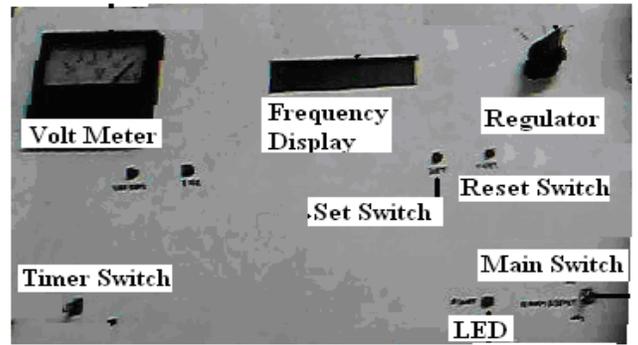
Anthropometric data

S.no	Age	Sex	Weight	Height (cms)	Hand Circumference		Hand Width		Wrist
					without thumb	with thumb	without thumb	with thumb	
1	30	M	75	165.1	21	25	10.5	11.5	17
2	32	M	65	170.18	20	26	9.5	12	18
3	50	M	65	175.26	22	27	11	13	20
4	52	M	60	167.64	21	24	10	11.5	16
5	48	M	55	175.26	20	23	9.5	10.5	15
6	49	M	80	160.02	22	25	11	12	18
7	40	M	60	167.64	20	23	9	10.5	17
8	32	M	60	154.94	23	27	11	13	19
9	36	M	70	177.8	22	26	10.5	13	17
10	59	M	90	180.34	23	27	13	15	21
11	25	F	50	162.56	18	21	8.5	10	15
12	56	F	85	139.7	19	22	8.5	10	17
13	29	F	80	149.86	20	22	9	11	16
14	20	F	55	134.62	18	21	8.5	10	16
15	29	F	57	162.56	19	21	9.5	12	16
16	48	F	65	177.8	20	22	10	11	15
17	18	F	55	177.8	19	21	9	11	16
18	28	F	60	152.4	18	22	8.5	10	15
19	27	F	62	165.1	20	23	9	12	17
20	19	F	65	167.64	21	20	8	10	16

Results & discussions

Analysis of data for males

s.no	Grip Strength in Kgf			
	Before experiment	After 1min	After 3 min	After 5min
m1	62	55	50	45
m2	70	62	56	50
m3	85	75	68	61
m4	60	52	47	42
m5	60	53	48	43
m6	62	54	49	44
m7	65	57	51	45
m8	70	62	56	50
m9	60	53	48	44
m10	80	71	63	57
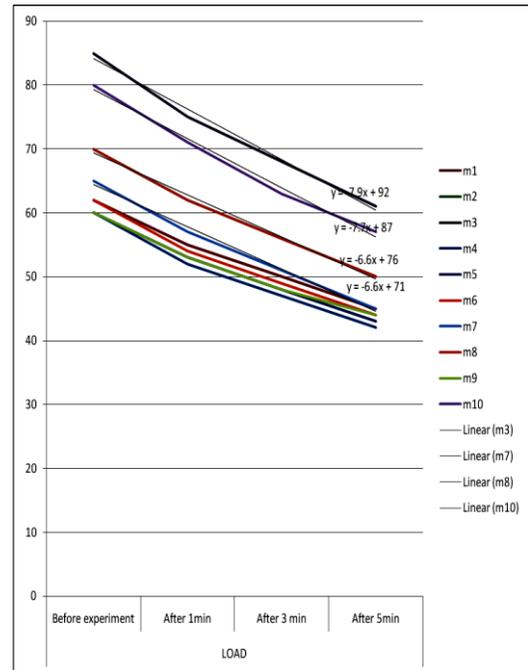
Analysis of data for females

S.no	Grip Strength in Kgf			
	Before experiment	After 1min	After 3 min	After 5min
F1	30	25	20	18
F2	15	15	10	12
F3	25	20	18	15
F4	45	39	35	30
F5	35	30	28	25
F6	30	26	23	20
F7	35	30	27	24
F8	30	26	23	20
F9	25	22	20	17
F10	40	35	31	27

General grip strength equation For Male

From the analysis we can generate the general line trend Equation for males i.e:

$$Y = -7.2 x + 81.5$$

General grip strength equation For Female
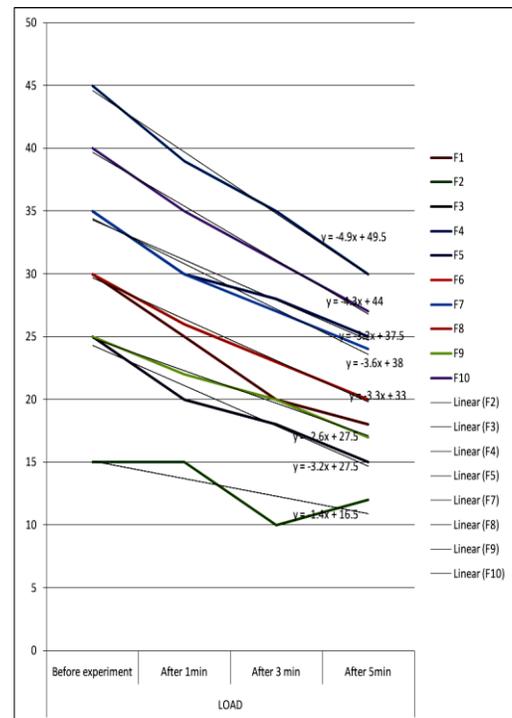
From the analysis we can generate the general line trend Equation for females **i.e:**

$$Y = -3.5x + 36.71$$

Effect of sex on grip strength
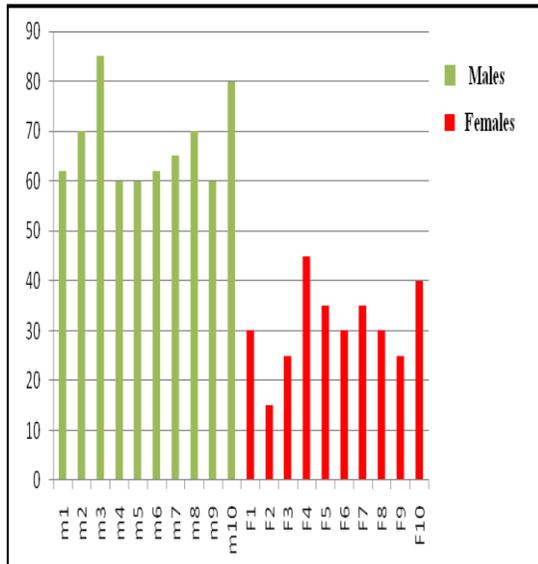
As per the experimented data it is seen that the average grip strength of male is 74 Kgf and that of female is 31 Kgf. This indicates that females have **54%** less grip strength than males.

Effect of weight on grip strength for males
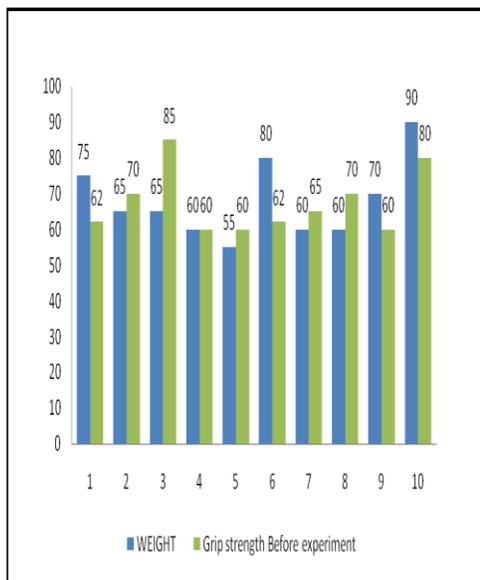
For males that their grip strength in general matches with their weight. For different males it can be seen that the value of their grip strength is near to the value of their weight.

Effect of weight on grip strength for females
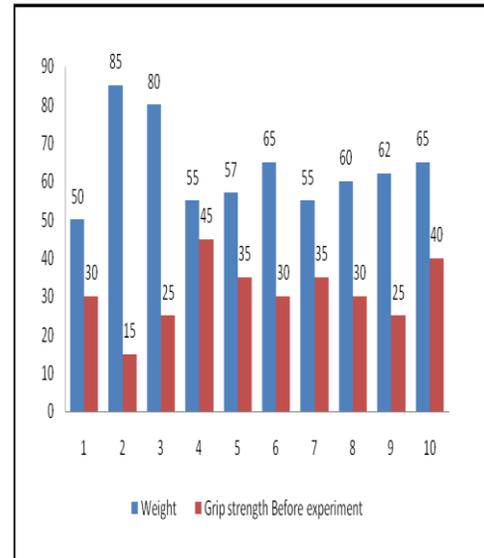
it can be seen that for overweight females (weight more than 55kg) there is a large difference in their grip strength. Or we can say for overweight females they have very low grip strength value as compare to average weight females. And about the average weight females they show the same trend as males i.e their grip strength value remains nearer to their weight taken in Kg

Standard deviation
Using the formula for standard deviation

$$\sigma 1 = V\sqrt{\sum\left(\frac{Dm - D\prime m}{n-1}\right)^2}$$

Between one minute and 3 minute data

for male is 2.06
for female is 1.16

Between three minute and five minute data

for male is 1.95
for female is 1.127

CONCLUSION
(a) General line trend equation for male considering all the trends comes to be

$$Y = -7.2X + 81.5$$

(b) General line trend equation for male considering all the trends comes to be

$$Y = -3.58\ X + 36.71$$

(c) Females have **54%** less grip strength force than males

(d) For males that their grip strength in general matches with their weight. For different males it can be seen that the value of their grip strength is near to the value of their weight

(e) For females it can be seen that for overweight females (weight more than 55kg) there is a large difference in their grip strength. Or we can say for overweight females they have very low grip strength value as compare to average weight females.

And about the average weight females they show the same trend as males i.e their grip strength value remains nearer to their weight taken in Kg

(f) Between one minute and three minute the standard deviation for male is 2.06
Between three minute and five minute the standard deviation for male is 1.95
Between one minute and three minute the standard deviation for female is 1.16
Between three minute and five minute the standard deviation for female is   1.127
Between three minute and five minute the standard deviation for female is   1.127

References

1. Muscle response to pneumatic hand tool torque reaction forces
   Robert   G.   Radwin; Ernst   Vanbergeijk; Thomas   J. Armstrong
   Ergonomics, 1366-5847, Volume 32, Issue 6, 1989, Pages 655 – 673
2. Effect of Grip Span on Lateral Pinch Grip Strength
   Carrie L. Shivers North Carolina State University, Raleigh, North Carolina
   Gary A. Mirka North Carolina State University, Raleigh, North Carolina
   David B. Kaber North Carolina State University, Raleigh, North Carolina
3. The effect of thumb interphalangeal joint position on strength of key pinch. Apfel,E.(1986).Journal of hand surjury,11A,47-51
4. A comparison of dominant and non dominant hand strengths.
   Armstrong, C.A.,& Oldham,J.A.(1999).Journal of hand surjury,24B,421-425.
5. Carpal tunnel syndrome and selected personal attributes.
   Armstrong,T.J.& Chaffin, D.B.(1979).Journal of occupational medicine.21.481-486.
6. The effects of instruction of finger strength measurements: Applicability of the cadwell regimen. Berg,V.J.,Clay,D.J.,Fathallah, F.A.,& Higginbotham,V.L.(1988). In F.Aghazadeh (Ed.).Trends in ergonomics/human factors V (pp.191-198).Amsterdam:Elsevier Science
7. A proposed standard procedure for static muscle strength testing. Caldwell,L.S.,Chaffin,D.B.,DukesDubos,F.N.,Kromer,K.H.E.,Laubacj,L.L.,Snook,S,H.,& Wasserman,D.E.(1974) Amrican industrial Hygiene Association journal.201-206.

Books—
1. Human factor ergonomics for building and construction by Martin halander
2. Ergonomics by Murrell K.H.F
3. Advances in industrial ergonomics and safety by Anil mital

AUTOMATED UNMANNED RAILWAY LEVEL CROSSING SYSTEM

J. BANUCHANDAR[#], V. KALIRAJ[#], P. BALASUBRAMANIAN[#], S. DEEPA[*], N. THAMILARASI[*]

[#]P.S.R ENGINEERING COLLEGE, STUDENT, B.E (ECE), SIVAKASI, INDIA
[*]ASSISTANT PROFESSOR, DEPARTMENT OF ECE, P.S.R ENGINEERING COLLEGE, SIVAKASI, INDIA

Abstract

In the rapidly flourishing country like ours, accidents in the unmanned level crossings are increasing day by day No fruitful steps have been taken so far in these areas. our paper deals with automatic railway gate operation (i.e.,) automatic railway gate at a level crossing replacing the gates operated by the gatekeepers, It deals with two things, Firstly it deals with the reduction of time for which the gate is being kept closed and secondly, to provide safety to the road users by reducing the accidents. By employing the automatic railway gate control at the level crossing the arrival of the train is detected by the sensors placed near to the gate. Hence, the time for which it is closed is less compared to the manually operated gates. The operation is automatic; error due to manual operation is prevented. Automatic railway gate control is highly microcontroller based arrangements, designed for use in almost all the unmanned level crossing in the train.

Keywords: Railway gate, level crossing.

1. INTRODUCTION

The place where track and highway/road intersects each other at the same level is known as "level crossing". There are mainly two types of level crossing they are Manned level crossing and Unmanned level crossing. Manned level crossing is classified into spl.Class, 'A'Class, 'B'Class, 'C'Class. Unmanned level crossing is classified into 'C'Class, 'D'Class. Railways being the cheapest mode of transportation are preferred over all the other means .When we go through the daily newspapers we come across many railway accidents occurring at unmanned railway crossings. This is mainly due to the carelessness in manual operations or lack of workers. We, in this paper have come up with a solution for the same. Using simple electronic components we have tried to automate the control of railway gates. As a train approaches the railway crossing from either side, the sensors placed at a certain distance from the gate detects the approaching train and accordingly controls the operation of the gate. When the wheels of the train moves over, both tracks are shorted to ground and this acts as a signal to the microcontroller indicating train arrival.

Also an indicator light has been provided to alert the motorists about the approaching train.

2. ACCIDENT AVOIDENCE DETAILS

When the train arrives in a particular direction the transmitter IR senses and generates appropriate signal, then at the same time the receiver IR receives the signal and generates an interrupt.

When the interrupt is generated the stepper motor rotates in clockwise direction. When the interrupt ends the stepper motor rotates in anti clock wise direction.

RAILWAY CROSSING
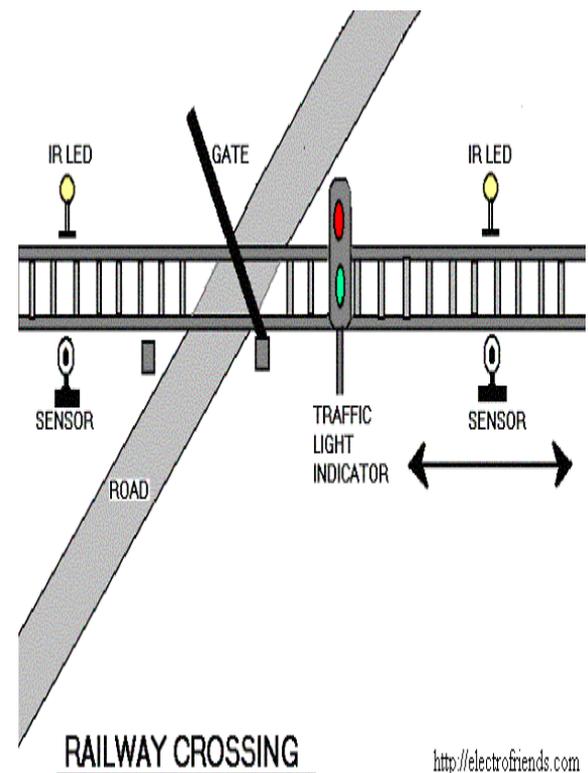
HARDWARE   IMPLEMENTATION

Micro Controller
Totally 40-pin DIP package manufactured with CMOS Technology.

L293D (motor driver)
Racially L293D 16DIP /ULN 2003 IC is used to drive the stepper motor.

STEPPER MOTOR
This is used to open and close the gates automatically when it is rotated clock wise or anticlockwise direction.
Stepper motor requires 500m amps current, so use the uln2003 or L293D drivers to drive the stepper motor.

SOFTWARE   IMPLEMENTATION

Keil software

3. BLOCK DIAGRAM DESCRIPTION
The block diagram consists of six major blocks, they are IR sensors, Microcontroller, L293D, Stepper motor, gate and power supply

3.1 IR SENSORS
Two IR sensor pairs (331,333) are used for transmitting and receiving signals.

3.1.1 IR CIRCUITS
This circuit has two stages: a transmitter unit and a receiver unit. The transmitter unit consists of an infrared LED and its associated circuitry.

3.1.2 IR TRANSMITTER
The transmitter circuit consists of the following components:

1. Resistors
2. IR LED

The IR LED emitting infrared light is put on in the transmitting unit. Infrared LED is driven through transistor BC 548.

Transmitter
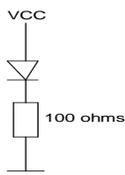
3.1.3 IR RECEIVER
The receiver circuit consists of the following components:

1. Resistors.
2. IR LED.

The receiver unit consists of a sensor and its associated circuitry. In receiver section, the first part is a sensor, which detects IR pulses transmitted by IR-LED. Whenever a train crosses the sensor, the output of IR sensor momentarily transits through a low state.
As a result the monostable is triggered and a short pulse is applied to the port pin of the 8051 microcontroller. On receiving a pulse from the sensor circuit, the controller activates the circuitry required for closing and opening of the gates and for track switching. The IR receiver circuit is shown in the figure below.

Receiver
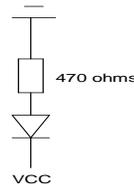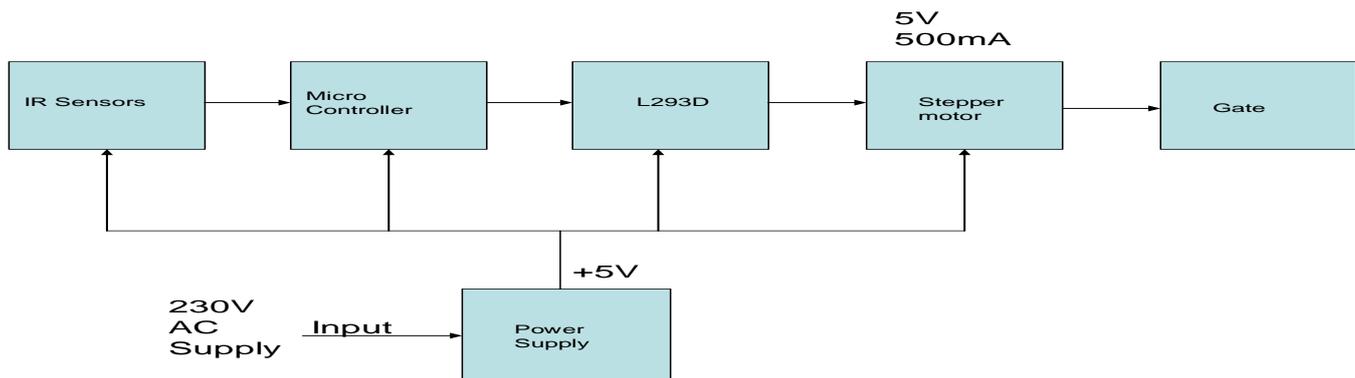
3.2 MICROCONTROLLER

It is designed using 8051 microcontroller to avoid railway accidents happening at unattended railway gates.

The Micro controller is a low power; high performance CMOS 8-bit micro controller with 4K bytes of Flash programmable and erasable read only memory (PEROM). The on-chip Flash allows the program memory to be reprogrammed in-system or by a conventional non-volatile memory programmer. By combining a versatile 8-bit CPU with Flash on a monolithic chip, the Atmel is a powerful microcomputer, which provides a highly flexible and cost-effective solution to many embedded control applications. By using this controller the data inputs from the smart card is passed to the Parallel Port of the pc and accordingly the software responds. The IDE for writing the embedded program used is KEI L software

FEATURES OF MICROCONTROLLER

The AT89C52 provides the following standard features: 8K bytes of Flash, 256 bytes of RAM, 32 I/O lines, three 16-bit timer/counters, a six-vector two-level interrupt architecture, a full duplex serial port, on-chip oscillator, and clock circuitry. In addition, the AT89C52 is designed with static logic for operation down to zero frequency and supports two software selectable power saving modes.
The Idle Mode stops the CPU while allowing the RAM, timer/counters, serial port, and interrupt system to continue functioning. The Power Down Mode saves the RAM contents but freezes the oscillator, disabling all other chip functions until the next hardware reset.

3.2.1 Keil Micro vision Integrated Development Environment.

Keil Software development tools for the 8051 micro controller family support every level of developer from the professional applications engineer to the student just learning about embedded software development. The industry-standard Keil C Compilers, Macro Assemblers, Debuggers, Real-time Kernels, and Single-board Computers support ALL 8051-compatible derivatives and help you get your projects completed on schedule.
The source code is written in assembly language .It is saved as ASM file with an extension. A51.the ASM file is converted into hex file using keil software. Hex file is dumped into micro controller using LABTOOL software. At once the file is dumped and the ROM is burnt then it becomes an embedded one.

3.3 L293D PUSH-PULL FOUR CHANNEL DRIVER WITH DIODES

➢ 600ma output current capability per channel
➢ 1.2a peak output current (non repetitive) per channel
➢ enable facility
➢ over temperature protection
➢ logical "0" input voltage up to 1.5 v
➢ (high noise immunity)
➢ internal clamp diodes

The Device is a monolithic integrated high voltage, high current four channel driver designed to accept standard DTL or TTL logic levels and drive inductive loads (such as relays solenoids, DC and stepping motors) and switching power transistors.

To simplify use as two bridges each pair of channels is equipped with an enable input. A separate supply input is provided for the logic, allowing operation at a lower voltage and internal clamp diodes are included.

This device is suitable for use in switching applications at frequencies up to 5 kHz. The L293D is assembled in a 16 lead plastic package which has 4 center pins connected together and used for heat sinking. The L293DD is assembled in a 20 lead surface mount which has 8 center pins connected together and used for heat sinking.

BLOCK DIAGRAM
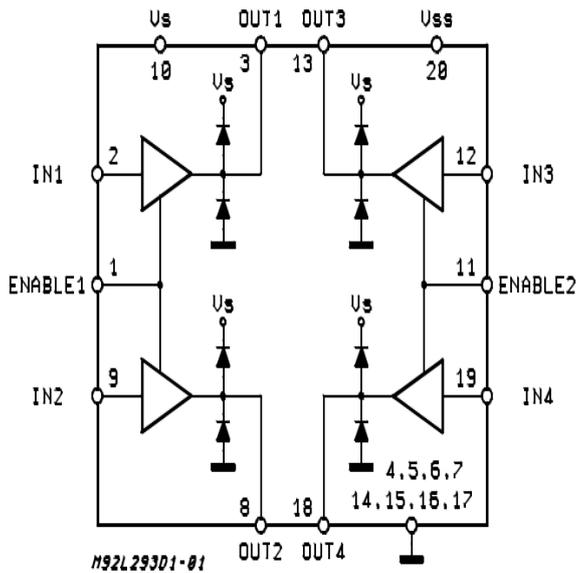
3.3.2 ADVANTAGES AND DISADVANTAGES OF L293D
ADVANTAGES

Efficient way of speed control of DC motor.
Produces more torque.
Produces less noise.

DISADVANTAGES

It is not applicable for AC motor.

APPLICATIONS

Industries.
Traction.
Home appliance.

3.4 STEPPER MOTOR

The stepper tutorial deals with the basic final stage drive circuitry for stepping motors. This circuitry is centered on a single issue, switching the current in each motor winding on and off, and controlling its direction. The circuitry discussed in this section is connected directly to the motor windings and the motor power supply, and this circuitry is controlled by a digital system that determines when the switches are turned on or off.

This section covers all types of motors, from the elementary circuitry needed to control a variable reluctance motor, to the H-bridge circuitry needed to control a bipolar permanent magnet motor. Each class of drive circuit is illustrated with practical examples, but these examples are not intended as an exhaustive catalog of the commercially available control circuits, nor is the information given here intended to substitute for the information found on the manufacturer's component data sheets for the parts mentioned.

This section only covers the most elementary control circuitry for each class of motor. All of these circuits assume that the motor power supply provides a drive voltage no greater than the motor's rated voltage, and this significantly limits motor performance. The next section, on current limited drive circuitry, covers practical high-performance drive circuits.

3.4.1 Stepping Sequences for a Four-Phase Unipolar Permanent Magnet Stepper Motor

This kind of motor has four coils which, when energized in the correct sequence, cause the permanent magnet attached to the shaft to rotate.

There are two basic step sequences. After step 4, the sequence is repeated from step 1 again.

Step	Coil 4	Coil 3	Coil 2	Coil 1	
a.1	on	off	off	Off	
a.2	off	on	off	Off	
a.3	off	off	on	Off	
a.4	off	off	off	On	

Reversing the order of the steps in a sequence will reverse the direction of rotation.

Here are some possible connection diagrams and some software

3.4.2 Single-Coil Excitation - Each successive coil is energized in turn.

This sequence produces the smoothest movement and consumes least power.

Step	Coil 4	Coil 3	Coil 2	Coil 1	
a.1	on	off	off	Off	
b.1	on	on	off	Off	
a.2	off	on	off	Off	
b.2	off	on	on	Off	
a.3	off	off	on	Off	
b.3	off	off	on	On	
a.4	off	off	off	On	
b.4	on	off	off	On	
b.4	on	off	off	On	

3.4.3 Two-Coil Excitation - Each successive pair of adjacent coils is energized in turn.

This is not as smooth and uses more power but produces greater torque

The excitation of Coil 1 is always the inverse of the excitation of Coil 3.

So, with the right circuit the excitation of Coil 4 is always the inverse of the excitation of Coil 2. You can generate this sequence with only two data lines. **Interleaving** the two sequences will cause the motor to **half-step.**

4. APPLICATIONS
➢ Real time transport systems.

5. ADVANTAGES
➢ Accident avoidance.
➢ Human Resource.
➢ Safety and quality of services

6. CONCLUSION
The accidents are avoided at places where there is no person managing the railway crossing gates.

Here we use the stepper motor to open and close the gates automatically when it is rotated clockwise or anticlockwise direction.

When the train arrives in a particular direction the transmitter IR senses and generates appropriate signal, then at the same time the receiver IR receives the signal and generates an interrupt.

When the interrupt is generated the stepper motor rotates in clockwise direction. When the interrupt ends the stepper motor rotates in anti clock wise direction

Railway Track
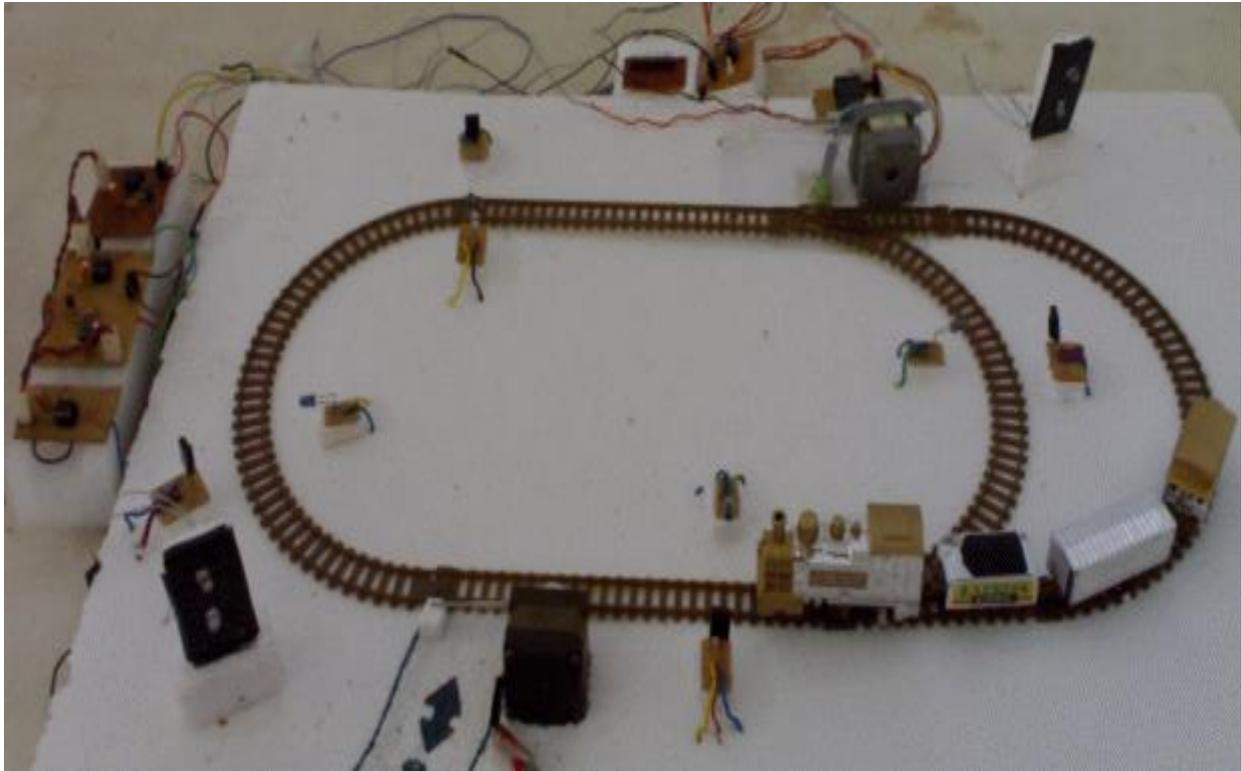
REFERENCES

➤ Adler, R. B., A. C. Smith, and R. L. Longani: "Introduction to Semiconductor Physics," vol. 1, p. 78, Semiconductor

➤ Electronics Education Comitee, John Wiley & Sons, Inc., New York ,1964.

➤ Schade, O. H.: "Analysis of Rectifier Operation", proc. IRE, vol.31, pp. 341-361, July, 1943.

➤ Stout, M. B.: "Analysis of Rectifier Circuits", Elec. Eng., vol. 54, September, 1935.

➤ Jacob Millman Christos C. Halkias.: "Electronic Devices And Circuits", Tata McGraw-Hill Publishing Company Ltd. Sep, 2003.

➤ The 8051 Microcontroller and Embedded Systems using Assembly and

➤ C by Muhammad Ali Mazidi, Janice Gillispie, Rolin D.Mckinlay.

➤ Part of stepping motors by Douglas W.Jones, the university of IOWA Department of computer science.

A mixed finite element approximation of the Stokes equations with the boundary condition of type (D+N)

Jaouad El-Mekkaoui[1], Abdeslam Elakkad[1,2], and Ahmed Elkhalfi[1]

[1] Laboratoire Génie Mécanique - Faculté des Sciences et Techniques B.P. 2202 - Route d'Imouzzer - Fès Maroc
[2] Discipline: Mathématiques, Centre de Formation des Instituteurs Sefrou, B.P: 243 Sefrou Maroc

ABSTRACT

In this paper we introduced the Stokes equations with a boundary condition of type (D+N).

The weak formulation obtained is a problem of saddle point type. We have shown the existence and uniqueness of the solution of this problem. We used the discretization by mixed finite element method with a posteriori error estimation of the computed solutions. In order to evaluate the performance of the method, the numerical results are compared with some previously published works or with others coming from commercial code like Adina system.

Keywords - **Stokes Equations, Mixed Finite element method, a posteriori error estimation, Adina system**.

I. INTRODUCTION

In modeling flow in porous media, it is essential to use a discretization method which satisfies the physics of the problem, i.e. conserve mass locally and preserve continuity of flux. The Raviart-Thomas Mixed Finite Element (MFE) method of lowest order satisfies these properties. Moreover, both the pressure and the velocity are approximated with the same order of convergence [4, 6]. The discretization of the velocity is based on the properties of Raviart-Thomas. Other works have been introduced by Brezzi, Fortin, Marini, Dougla and Robert [4, 5, 7].

This method was widely used for the prediction of the behavior of fluid in the hydrocarbons tank.
 A posteriori error analysis in problems related to fluid dynamics is a subject that has received a lot of attention during the last decades. In the conforming case there are several ways to define error estimators by using the residual equation. in particular, for the Stokes problem, M. Ainsworth, J. Oden [9], R.E. Bank, B.D. Welfert [10], C. Crestensen, S.A. Funken [11], D. Kay, D. Silvester [12] and R. Verfurth [13], introduced several error estimators and provided that that they are equivalent to the energy norm of the errors. Other works for the stationary Navier-Stokes problem have been introduced in [14, 17, 18, 20, 16].

This paper describes a numerical solution of Stokes equations with a boundary condition noted (D+N). For the equations, we offer a choice of tow-dimensional domains on which the problem can be posed, along with boundary conditions and other aspects of the problem, and a choice of finite element discretization on a rectangular element mesh. The plan of the paper is as follows. The model problem is

described in sections II, In Section III, we prove the existence and uniqueness of the solution of the weak formulation obtained, followed by a mixed finite element discretization for the Stokes equations in section IV. In section V we consider a posteriori error bounds of the computed solution, and numerical experiments are carried out in section VI.

II. GOVERNING EQUATIONS

We consider the Stokes equations for the flow;

$$-\nabla^2 \vec{u} + \nabla p = \vec{f} \qquad in \ \Omega \tag{1}$$

$$\nabla.\vec{u} = 0 \qquad in \ \Omega \tag{2}$$

$$(D+N): \qquad b_0 \vec{u} + \frac{\partial \vec{u}}{\partial n} - \vec{n}p = \vec{t} \quad in \ \Gamma \tag{3}$$

$\vec{u}$ is the fluid velocity, p is the pressure field. ∇ is the gradient, $\nabla.$ is the divergence and ∇^2 is the laplacien operator, $\vec{f} \in [L^2(\Omega)]^2$. Ω is a bounded and connected domain of $\mathbb{R}^2$ with a Lipchitz continuous boundary $\Gamma = \partial\Omega$. where $\vec{n}$ denote the outward pointing normal to the boundary, and

$\vec{t} \in [L^2(\Gamma)]^2$. b_0 is a function defined and bounded on Γ verify:

$$\exists \ \alpha_0 > 0 \ such \ that \ b_0 \geq \alpha_0 \ almost \ everywhere. \tag{4}$$

We define the spaces:

$$h^1(\Omega) = \left\{ u: \Omega \rightarrow \mathbb{R}/u \ ; \ \frac{\partial u}{\partial x} \ ; \ \frac{\partial u}{\partial y} \in L^2(\Omega) \right\} \tag{5}$$

$$H^1(\Omega) = [h^1(\Omega)]^2 \tag{6}$$

$$h^2(\Omega) = \left\{ u: \Omega \rightarrow \mathbb{R}/u ; \ \frac{\partial u}{\partial x_i} ; \ \frac{\partial^2 u}{\partial x_i \partial x_j} \in L^2(\Omega); \ i,j = 1,2 \right\} \tag{7}$$

Then the standard weak formulation of the Stokes flow problem (1)-(2)-(3) is the following:

$$\begin{cases} find\ (\vec{u},p) \in H^1(\Omega) \times L^2(\Omega)\ such\ that: \\ \int_\Omega \nabla\vec{u}:\nabla\vec{v}\ d\Omega + \int_\Gamma b_0\,\vec{u}.\vec{v}\ d\gamma - \int_\Omega p\nabla.\vec{v}\ d\Omega = \int_\Gamma \vec{t}.\vec{v}\ d\gamma + \int_\Omega \vec{f}.\vec{v}\ d\Omega\ for\ all\ \vec{v} \in H^1(\Omega) \\ -\int_\Omega q\nabla.\vec{u}\ d\Omega = 0\ \ for\ all\ q \in L^2(\Omega) \end{cases}$$

$$(8)$$

Let the bilinear forms $a: H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ and $b: H^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$

$$a(\vec{u},\vec{v}) = \int_\Omega \nabla\vec{u}:\nabla\vec{v}\ d\Omega + \int_\Gamma b_0\,\vec{u}.\vec{v}\ d\gamma \qquad (9)$$

$$b(\vec{u},q) = -\int_\Omega q\nabla.\vec{u}\ d\Omega. \qquad (10)$$

Given the functional $l: H^1(\Omega) \to \mathbb{R}$,

$$l(\vec{v}) = \int_\Gamma \vec{t}.\vec{v}\ d\gamma + \int_\Omega \vec{f}.\vec{v}\,d\Omega. \qquad (11)$$

The underlying weak formulation (8) may be restated as

$$\begin{cases} find\ (\vec{u},p) \in H^1(\Omega) \times L^2(\Omega)\ such\ that: \\ a(\vec{u},\vec{v}) + b(\vec{u},q) = l(\vec{v})\quad for\ all\ \in \vec{v}\ H^1(\Omega) \\ b(\vec{u},q) = 0\quad for\ all\ q \in L^2(\Omega) \end{cases}$$

$$(12)$$

III. THE EXISTENCE AND UNIQUENESS OF THE SOLUTION

In this section we will show that the problem (12) has exactly one solution $(\vec{u},p) \in H^1(\Omega) \times L^2(\Omega)$.

It suffices to verify that the bilinear form a is positive, continuous and $H^1(\Omega) - elliptic$, and the bilinear form b is continuous and satisfies the inf-sup condition (see theorem 6.8 in [1]).

Theorem 3.1. $H^1(\Omega)$ is a real Hilbert space, with norm denoted by $\|.\|_{J,\Omega}$, for the scalar product:

$$\langle \vec{u},\vec{v}\rangle_J = \int_\Omega \nabla\vec{u}:\nabla\vec{v}\ d\Omega + \int_\Gamma \vec{u}.\vec{v}\ d\gamma \qquad (13)$$

$$\|\vec{u}\|_{J,\Omega} = \langle \vec{u},\vec{u}\rangle_J^{\frac{1}{2}} = \left(\|\nabla\vec{u}\|_{0,\Omega}{}^2 + \|\vec{u}\|_{0,\Gamma}{}^2 \right)^{\frac{1}{2}}. \qquad (14)$$

To prove this theorem we need the following lemma.

Lemma 3.2. There are two strictly positive constants c_1 and c_2 such that:

$$c_1\|\vec{v}\|_{1,\Omega} \leq \|\vec{v}\|_{J,\Omega} \leq c_2\|\vec{v}\|_{1,\Omega}\ for\ all\ \vec{v} \in H^1(\Omega). \qquad (15)$$

Proof. The mapping $\gamma_0: H^1(\Omega) \to L^2(\Gamma)$ is continuous (see the theorem 1.2 in [21]), then there exists $c > 0\ such\ that:$

$$\|\vec{v}\|_{0,\Gamma} \leq c\|\vec{v}\|_{1,\Omega}\ for\ all\ \vec{v} \in H^1(\Omega)$$

We have also: $\|\nabla\vec{v}\|_{0,\Omega} \leq \|\vec{v}\|_{1,\Omega}\ for\ all\ \vec{v} \in H^1(\Omega)$, then

$$\|\vec{v}\|_{J,\Omega} = \left(\|\nabla\vec{v}\|_{0,\Omega}{}^2 + \|\vec{v}\|_{0,\Gamma}{}^2 \right)^{\frac{1}{2}} \leq c_2\|\vec{v}\|_{1,\Omega}$$

$for\ all\ \vec{v} \in H^1(\Omega)$, with $c_2 = (c^2+1)^{\frac{1}{2}}$

On the other hand, there exists a constant $l > 0\ such\ that$

$$\|\vec{v}\|_{0,\Omega} \leq l(\|\nabla\vec{v}\|_{0,\Omega} + \|\vec{v}\|_{0,\Gamma})\ for\ all\ \vec{v} \in H^1(\Omega)$$

(see 5.55 in [1])

i.e.

$$\|\vec{v}\|_{0,\Omega}^2 \leq l^2\left(\|\nabla\vec{v}\|_{0,\Omega} + \|\vec{v}\|_{0,\Gamma}\right)^2 \leq 2l^2\left(\|\nabla\vec{v}\|_{0,\Omega}^2 + \|\vec{v}\|_{0,\Gamma}^2\right) = 2l^2\|\vec{v}\|_{J,\Omega}$$

We have also: $\|\nabla\vec{v}\|_{0,\Omega} \leq \|\vec{v}\|_{1,\Omega}\ for\ all\ \vec{v} \in H^1(\Omega)$.

Then, $c_1\|\vec{v}\|_{1,\Omega} \leq \|\vec{v}\|_{J,\Omega}$ for all $\vec{v} \in H^1(\Omega)$, with

$$c_1 = \left(\frac{1}{2l^2+1}\right)^{\frac{1}{2}}.$$

Proof of theorem 3.1. Ω is a bounded and connected domain of $\mathbb{R}^2$, then it is easy to verify that $\langle.,.\rangle_J$ is a scalar product, i.e. $\left(H^1(\Omega), \|.\|_{J,\Omega}\right)$ is an Euclidean space.

We have $\left(H^1(\Omega), \|.\|_{1,\Omega}\right)$ is a real Hilbert, then it is complete. By the lemma 3.2, $\|.\|_{1,\Omega}$ and $\|.\|_{J,\Omega}$ are two equivalents norms, then $\left(H^1(\Omega), \|.\|_{J,\Omega}\right)$ is complete, therefore it is a real Hilbert space.

Theorem 3.3. $i)\ a$ is continuous.

$ii)\ a$ is $H^1(\Omega) - elliptic$ for the norm $\|.\|_{J,\Omega}$, i.e. there exists a constant $\alpha > 0\ such\ that$

$$a(\vec{v},\vec{v}) \geq \alpha\|\vec{v}\|_{J,\Omega}^2\ for\ all\ \vec{v} \in H^1(\Omega). \qquad (16)$$

Proof. $i)\ a$ is a scalar product, by Cauchy-Schwarz inequality we have

$$a(\vec{u},\vec{v}) \le \left(a(\vec{u},\vec{u}) \times a(\vec{v},\vec{v})\right)^{\frac{1}{2}} \le M \|\vec{u}\|_{J,\Omega}. \|\vec{v}\|_{J,\Omega}$$
$$for\ all\ \vec{u},\ \vec{v} \in H^1(\Omega) \tag{17}$$

with $M = Max(1;\ \sup_{x \in \Gamma} b_0(x))$.

ii) Let $\vec{v} \in H^1(\Omega)$, using (4) then gives:

$$a(\vec{v},\vec{v}) = \int_\Omega (\nabla \vec{v})^2 \,d\Omega + \int_\Gamma b_0(\vec{v})^2 \,d\gamma \ge \alpha \|\vec{v}\|_{J,\Omega}^2,$$

With $\alpha = \min(1; \alpha_0)$.

Theorem 3.4. $i)\ b\ is\ continuous.$

         $ii)$ b is satisfies the inf-sup: There exists a constant $\beta > 0$ such that

$$\inf_{0 \ne q \in L^2(\Omega)} \sup_{0 \ne \vec{v} \in H^1(\Omega)} \frac{\int_\Omega q \nabla.\vec{v} \,d\Omega}{\|q\|_{0,\Omega} \|\vec{v}\|_{J,\Omega}} \ge \beta. \tag{18}$$

To prove this theorem we need the following lemmas:
Lemma 3.5.

$$\nabla.:\ \left(H^1(\Omega), \|.\|_{J,\Omega}\right) \to \left(L^2(\Omega), \|.\|_{0,\Omega}\right)$$
$$\vec{v} \mapsto \nabla.\vec{v}$$

is continuous linear mapping and $R(\nabla.) = L^2(\Omega)$.

Proof. It is clear that $\nabla.$ is a linear mapping. Remains to chow that it is continuous and
$R(\nabla.) = L^2(\Omega)$.
We have: $\|\nabla.\vec{v}\|_{0,\Omega} \le \sqrt{2}\|\nabla \vec{v}\|_{0,\Omega} \le \sqrt{2}\|\vec{v}\|_{J,\Omega}$ for all
$\vec{v} \in H^1(\Omega)$, then $\nabla.$ is continuous
    Let $q \in L^2(\Omega)$, then $\int_\Omega q \,d\Omega \in \mathbb{R}$.
We set $k = \int_\Omega q \,d\Omega$,
then $q - \frac{k}{|\Omega|} \in L_0^2(\Omega) = \{f \in L^2(\Omega); \int_\Omega f \,d\Omega = 0\}$,
with $|\Omega|$ is the area of Ω.
Since $L_0^2(\Omega)$ is the range space of the linear mapping $\nabla.$:
$H_0^1(\Omega) \to L_0^2(\Omega)$ (see lemma 6.8 in $[1]$), then there exists
$\vec{u} \in H_0^1(\Omega)$ such that $\nabla.\vec{u} = q$.
Let $\vec{u}_0 = \left(\frac{k}{|\Omega|}x; 0\right)$ for all $(x,y) \in \mathbb{R}^2$, we have
$\nabla.(\vec{u} + \vec{u}_0) = q,$ where $R(\nabla.) = L^2(\Omega)$.

Lemma 3.6. There exists a constant $\beta > 0$ such that: for all $q \in L^2(\Omega)$ there exists $\vec{v} \in H^1(\Omega)$ such that $\nabla.\vec{v} = q$ and $\beta\|\vec{v}\|_{J,\Omega} \le \|q\|_{0,\Omega}$.

Proof. By the lemma 3.5, $R(\nabla.) = L^2(\Omega)$, then $R(\nabla.)$ is closed in $L^2(\Omega)$, therefore there exists $\beta > 0$ such that , for all $q \in L^2(\Omega)$ there exists $\vec{v} \in H^1(\Omega)$ such that $\nabla.\vec{v} = q$ and $\beta\|\vec{v}\|_{J,\Omega} \le \|q\|_{0,\Omega}$ (see the lemma A.3 [1]).

Proof of theoreom 3.4. i) Let $(\vec{v}, q) \in H^1(\Omega) \times L^2(\Omega)$; we have
$$b(\vec{v},q) = -\int_\Omega q\nabla.\vec{v} \,d\Omega \le \|q\|_{0,\Omega}\|\nabla.\vec{v}\|_{0,\Omega} \le \sqrt{2}\|q\|_{0,\Omega}\|\nabla\vec{v}\|_{0,\Omega} \le \sqrt{2}\|q\|_{0,\Omega}\|\vec{v}\|_{J,\Omega}$$
Then b is continuous.
ii) Let $q \in L^2(\Omega) - \{0\}$, by lemma 3.6, there exists $\vec{v'} \in H^1(\Omega)$ such that $\nabla.v' = q$ and
$\beta\|\vec{v'}\|_{J,\Omega} \le \|q\|_{0,\Omega}$. Then
$$\sup_{0 \ne \vec{v} \in H^1(\Omega)} \frac{\int_\Omega q\nabla.\vec{v} \,d\Omega}{\|\vec{v}\|_{J,\Omega}} \ge \frac{\int_\Omega q\nabla.\vec{v'} \,d\Omega}{\|\vec{v'}\|_{J,\Omega}} \ge \frac{\int_\Omega q^2 \,d\Omega}{\|\vec{v'}\|_{J,\Omega}} = \frac{\|q\|_{0,\Omega}^2}{\|\vec{v'}\|_{J,\Omega}} \ge \beta.\|q\|_{0,\Omega}$$

We define the "big" symmetric bilinear form
$$B\left((\vec{u},p);(\vec{v},q)\right) = a(\vec{u},\vec{v}) + b(\vec{v},p) + b(\vec{u},q) \tag{19}$$
And the corresponding function $F\left((\vec{v},q)\right) = l(\vec{v})$,
choosing the successive test vectors $(\vec{v},0)$ and $(\vec{0},q)$ shows that the stokes problem (12) can be rewritten in the form:
$$find\ (\vec{u},p) \in H^1(\Omega) \times L^2(\Omega)\ such\ that$$

$$B\left((\vec{u},p);(\vec{v},q)\right) = F\left((\vec{v},q)\right) for\ all\ (\vec{v},q) \in H^1(\Omega) \times L^2(\Omega). \tag{20}$$
The bilinear form is positive continuous and $H^1(\Omega) - elliptic,$ and the bilinear form b is continuous and satisfies the inf-sup condition. Then the problem (12) is well-posed and the ''B-stability bound'' [1], given below:

Proposition 3.7. $for\ all\ (\vec{w},s) \in H^1(\Omega) \times L^2(\Omega)$, we have that:
$$\sup_{(\vec{v},q) \in H^1(\Omega) \times L^2(\Omega)} \frac{B\left((\vec{w},s);(\vec{v},q)\right)}{\|\vec{v}\|_{J,\Omega} + \|q\|_{0,\Omega}} \ge \gamma_D \left(\|\vec{w}\|_{J,\Omega} + \|s\|_{0,\Omega}\right) \tag{21}$$
where γ_D depends only on the shape of the domain.

    The bilinear form a is symmetric, and continuous and semi positive definite on $H^1(\Omega)$, in this case we say the problem (12) is a type of saddle-point problem.
The theorem (3.3) and (3.4) ensure the existence and uniqueness of the solution of the problem (12) (see the theorem 6. 2 in [1]). In the following section we will solve this problem by mixed finite element method.

IV. MIXED FINITE ELEMENT APPROXIMATION
A discrete weak formulation is defined using finite dimensional spaces $X_1^h \subset H^1(\Omega)$ and $M^h \subset L_2(\Omega)$.

The discrete version of (12) is:

$$\begin{cases} find \ \overrightarrow{u_h} \ \epsilon X_1^h \ and \ p_h \ \epsilon M^h \ such \ that: \\ a(\overrightarrow{u_h}, \overrightarrow{v_h}) + b(\overrightarrow{v_h}, p_h) = l(\overrightarrow{v_h}) \ "for \ all" \ \overrightarrow{v_h} \epsilon \ X_1^h \quad (22) \\ b(\overrightarrow{u_h}, q_h) = 0 \quad\quad\quad\quad "for \ all" \ q_h \epsilon M^h \end{cases}$$

We use a set of vector-valued basis functions $\{\overrightarrow{\varphi_j}\}_{j=1,...,n_u}$,

so that

$$\overrightarrow{u_h} = \sum_{j=1}^{n_u} u_j \overrightarrow{\varphi_j} \tag{23}$$

We introduce a set of pressure basis functions $\{\psi_k\}_{k=1,...,n_p}$ and set

$$p_h = \sum_{k=1}^{n_p} p_k \psi_k \tag{24}$$

Where n_u and n_p are the numbers of velocity and pressure basis functions, respectively.

We find that the discrete formulation (22) can be expressed as a system of linear equations

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{25}$$

With

$$A = [a_{ij}]; \quad a_{ij} = \int_\Omega \nabla\overrightarrow{\varphi_i} : \nabla\overrightarrow{\varphi_j} d\Omega + \int_\Gamma b_0 \overrightarrow{\varphi_i} . \overrightarrow{\varphi_j} d\gamma \tag{26}$$

$$B = [b_{kj}]; \quad b_{kj} = -\int_\Omega \psi_k \nabla . \overrightarrow{\varphi_j} d\Omega \tag{27}$$

for $i, j = 1, ..., n_u$ and $k = 1, ..., n_p$.

The right-hand side vectors in (4.4) are

$$f = [f_i]; \quad f_i = \int_\Omega \overrightarrow{f} \overrightarrow{\varphi_i} d\Omega + \int_\Gamma \overrightarrow{t_i} . \overrightarrow{\varphi_i} d\gamma \tag{28}$$

for $i = 1, ..., n_u$,

and the function pair $(\overrightarrow{u}_h, p_h)$ obtained by substituting the solution vectors $u \in \mathbb{R}^{n_u}$ and $p \in \mathbb{R}^{n_p}$ into (23) and (24) is the mixed finite element solution. The system (25)-(28) is henceforth referred to as the discrete stokes problem.

We use the iterative methods Minimum Residual Method (MINRES) for solving the symmetric system.

V. A RESIDUAL ERROR ESTIMATOR

In this section we assume that $\overrightarrow{f}$ and $\overrightarrow{t}$ are the polynomials.

Let T_h ; $h > 0$, be a family of rectangulations of Ω. For any $T \epsilon T_h$, ω_T is of rectangles sharing at least one edge with element T, $\widetilde{\omega}_T$ is the set of rectangles sharing at least one vertex with T. Also, for an element edge E, ω_E denotes the union of rectangles sharing E, while $\widetilde{\omega}_E$ is the set of rectangles sharing at least one vertex whit E.

Next, ∂T is the set of the four edges of T we denote by $\varepsilon(T)$ and N_T the set of its edges and vertices, respectively.

We let $\varepsilon_h = \bigcup_{T \epsilon T_h} \varepsilon(T)$ denotes the set of all edges split into interior and boundary edges.

$$\varepsilon_h = \varepsilon_{h,\Omega} \cup \varepsilon_{h,\Gamma},$$

where
$$\varepsilon_{h,\Omega} = \{E \epsilon \varepsilon_h : E \subset \Omega\},$$
$$\varepsilon_{h,\Gamma} = \{E \epsilon \varepsilon_h : E \subset \partial\Omega\}.$$

The bubble functions on the reference element $\tilde{T} = (0,1) \times (0,1)$ are defined as follows:

$$b_{\tilde{T}} = 2^4 x(1-x)y(1-y)$$
$$b_{\tilde{E}_1, \tilde{T}} = 2^2 x(1-x)(1-y)$$
$$b_{\tilde{E}_2, \tilde{T}} = 2^2 y(1-y)x$$
$$b_{\tilde{E}_3, \tilde{T}} = 2^2 y(1-x)$$
$$b_{\tilde{E}_4, \tilde{T}} = 2^2 y(1-y)(1-x)$$

Here $b_{\tilde{T}}$ is the reference element bubble function, and $b_{\tilde{E}_i, \tilde{T}}$, $i = 1:4$ are reference edge bubble functions. For any $T \epsilon T_h$, the element bubble functions is $b_T = b_{\tilde{T}} o F_T$ and the element edge bubble function is $b_{E_i, T} = b_{\tilde{E}_i, \tilde{T}} o F_T$ where F_T the affine map form $\tilde{T}$ to T.

For an interior edge $E \epsilon \varepsilon_{h,\Omega}$, b_E is defined piecewise, so that $b_{E/T_i} = b_{E,T_i}$, $i = 1:2$, where $E = \overline{T}_1 \cap \overline{T}_2$. For a boundary edge $E \epsilon \varepsilon_{h,\Gamma}$, $b_E = b_{E,T}$, where T is the rectangle such that $E \epsilon \partial T$.

With these bubble functions, ceruse et al ([19], lemma 4.1] established the following lemma.

Lemma 5.1. Let T be an arbitrary rectangle in T_h and $E \in \partial T$.
For any $\overrightarrow{v}_T \epsilon P_{k_0}(T)$ and $\overrightarrow{v}_E \epsilon P_{k_1}(E)$, the following inequalities hold.

$$c_k \|\overrightarrow{v}_T\|_{0,T} \leq \left\| \overrightarrow{v}_T b_T^{\frac{1}{2}} \right\|_{0,T} \leq C_k \|\overrightarrow{v}_T\|_{0,T} \tag{29}$$

$$|\overrightarrow{v}_T b_T|_{1,T} \leq C_k h_T^{-1} \|\overrightarrow{v}_T\|_{0,T} \tag{30}$$

$$c_k \|\overrightarrow{v}_E\|_{0,E} \leq \left\| \overrightarrow{v}_E b_E^{\frac{1}{2}} \right\|_{0,E} \leq C_k \|\overrightarrow{v}_E\|_{0,E} \tag{31}$$

$$\|\overrightarrow{v}_E b_E\|_{0,T} \leq C_k h_E^{\frac{1}{2}} \|\overrightarrow{v}_E\|_{0,E} \tag{32}$$

$$|\overrightarrow{v}_E b_E|_{1,T} \leq C_k h_E^{-\frac{1}{2}} \|\overrightarrow{v}_E\|_{0,E}, \tag{33}$$

where c_k and C_k are tow constants which only depend on the element aspect ratio and the polynomial degrees k_0 and k_1.

Here, k_0 and k_1 are fixed and c_k and C_k can be associated with generic constants $c \ and \ C$. In addition, $\overrightarrow{v}_E$

which is only defined on the edge E also denotes its natural extension to the element T.

From the inequalities (32) and (33) we established the following lemma:

Lemma 5.2. Let T be an rectangle $and\ E \in \partial T \cap \varepsilon_{h,\Gamma}$. For any $\vec{v}_E \epsilon P_{k_1}(E)$, the following inequalities hold.

$$\|\vec{v}_E b_E\|_{J,T} \leq Ch_E^{-\frac{1}{2}}\|\vec{v}_E\|_{0,E}. \tag{34}$$

Proof. Since $\vec{v}_E b_E = \vec{0}$ in the other three edges of rectangle T, it can be extended to the whole of Ω by setting $\vec{v}_E b_E = \vec{0}$ in $\Omega\backslash\bar{T}$, then
$$\|\vec{v}_E b_E\|_{1,T} = \|\vec{v}_E b_E\|_{1,\Omega}$$
$$and\ \|\vec{v}_E b_E\|_{J,T} = \|\vec{v}_E b_E\|_{J,\Omega}.$$
Using the inequalities (32), (33) and the lemma (3.2) gives
$$\|\vec{v}_E b_E\|_{J,T} = \|\vec{v}_E b_E\|_{J,\Omega} \leq c_2\|\vec{v}_E b_E\|_{1,\Omega}$$
$$= c_2\|\vec{v}_E b_E\|_{1,T}$$
$$= c_2\left(\|\vec{v}_E b_E\|_{0,T}^2 + |\vec{v}_E b_E|_{1,T}^2\right)^{\frac{1}{2}}$$

$$\|\vec{v}_E b_E\|_{J,T} \leq c_2 C_k (h_E + h_E^{-1})^{\frac{1}{2}}\|\vec{v}_E\|_{0,E}$$

$$\leq c_2 C_k (D^2+1)^{\frac{1}{2}} h_E^{-\frac{1}{2}}\|\vec{v}_E\|_{0,E}$$

$$\leq Ch_E^{-\frac{1}{2}}\|\vec{v}_E\|_{0,E},,$$

where D is the diameter of Ω and $C = c_2 C_k (D^2+1)^{\frac{1}{2}}$.

We recall some quasi-interpolation estimates in the following lemma.

Lemma 5.3. Clement interpolation estimate:
Given $\vec{v} \epsilon H^1(\Omega)$, let $\vec{v}_h \epsilon X_h$ be the quasi-interpolant of $\vec{v}$ defined by averaging as in [20].
For any $T\epsilon T_h$, $\|\vec{v} - \vec{v}_h\|_{0,T} \leq Ch_T|\vec{v}|_{1,\tilde{\omega}_T}$, (35)

and for all $E\epsilon\partial T$, $\|\vec{v} - \vec{v}_h\|_{0,E} \leq Ch_E^{\frac{1}{2}}|\vec{v}|_{1,\tilde{\omega}_E}$. (36)

We let $(\bar{u},p)$ denote the solution of (12) and let $(\bar{u}_h,p_h)$ denote the solution of (22) with an approximation on a rectangular subdivision T_h.
Our aim is to estimate the velocity and the pressure errors $\vec{e} = \bar{u} - \bar{u}_h; \varepsilon = p - p_h$.
The element contribution $\eta_{R,T}$ of the residual error estimator η_R is given by

$$\eta_{R,T}^2 = h_T^2 \|\vec{R}_T\|_{0,T}^2 + \|R_T\|_{0,T}^2 + \sum_{E\in\partial T} h_E \|\vec{R}_E\|_{0,E}^2 \tag{37}$$

and the components in (37) are given by
$$\vec{R}_T = \{\vec{f} + \nabla^2\vec{u}_h - \nabla p_h\}_{/T} \tag{38}$$
$$R_T = \{\nabla.\vec{u}_h\}_{/T} \tag{39}$$

$$\vec{R}_E = \begin{cases} \frac{1}{2}[\![\nabla\vec{u}_h - p_h I]\!] \ ; E\epsilon\varepsilon_{h,\Omega} \\ \vec{t} - \left(b_0\vec{u}_h + \frac{\partial\vec{u}_h}{\partial\vec{n}_{E,T}} - p_h\vec{n}_{E,T}\right); E\epsilon\varepsilon_{h,\Gamma} \end{cases} \tag{40}$$

With the key contribution coming from the stress jump associated with an edge E adjoining elements T and S:

$$[\![\nabla\vec{u}_h - p_h I]\!] = \left((\nabla\vec{u}_h - p_h I)_{/T} - (\nabla\vec{u}_h - p_h I)_{/S}\right)\vec{n}_{E,T}$$

The global residual error estimator is given by:

$$\eta_R = \sqrt{\sum_{T\epsilon T_h} \eta_{R,T}^2}.$$

For any $T\epsilon T_h$, and $E \in \partial T$, we define the following two functions:
$$\vec{w}_T = \vec{R}_T b_T \ ; \quad \vec{w}_E = \vec{R}_E b_E.$$
$Since\ \vec{w}_T = \vec{0}$ on $\partial T, it$ can be extended to the whole of Ω by setting $\vec{w}_T = \vec{0}$ in $\Omega\backslash T$.
$- if\ E\epsilon\partial T \cap \varepsilon_{h,\Omega}\ then\ \vec{w}_E = \vec{0}$ in $\partial\omega_E$.
$- if\ E\epsilon\partial T \cap \varepsilon_{h,\Gamma}\ then\ \vec{w}_E = \vec{0}$ in the other three edges of rectangle T.

With these two functions we have the following lemmas:

Lemma 5.4. for any $T\epsilon T_h$ we have:
$$\int_T(\nabla\vec{u} - pI):\nabla\vec{w}_T d\Omega = \int_\Omega(\nabla\vec{u} - pI):\nabla\vec{w}_T d\Omega = \int_T\vec{f}.\vec{w}_T d\Omega \tag{41}$$
Proof.

$$\int_T(\nabla\vec{u} - pI):\nabla\vec{w}_T d\Omega = \int_{\partial T}(\nabla\vec{u} - pI)\vec{n}.\vec{w}_T d\gamma - \int_T(\nabla^2\vec{u} - \nabla p).\vec{w}_T d\Omega$$

Since $-\nabla^2\vec{u} + \nabla p = \vec{f}$ in Ω and $\vec{w}_T = \vec{0}$ in $\Omega\backslash T$

then:

$$\int_T(\nabla\vec{u} - pI):\nabla\vec{w}_T d\Omega = \int_\Omega(\nabla\vec{u} - pI):\nabla\vec{w}_T d\Omega = \int_T\vec{f}.\vec{w}_T d\Omega$$

Lemma 5.5. $i)\ if\ E\epsilon\partial T \cap \varepsilon_{h,\Omega}, we\ have:$
$$\int_{\omega_E}(\nabla\vec{u} - pI):\nabla\vec{w}_E d\Omega = \int_{\omega_E}\vec{f}.\vec{w}_E d\Omega. \tag{42}$$

$ii) if\ E\epsilon\partial T \cap \varepsilon_{h,\Gamma}, we\ have:$
$$\int_T(\nabla\vec{u} - pI):\nabla\vec{w}_E d\Omega = \int_{\partial T}(\vec{t} - b_0\vec{u})\vec{w}_E d\gamma + \int_T\vec{f}.\vec{w}_E d\Omega \tag{43}$$

Proof. $i)$ The same proof of (41).

$ii)$ *if* $E \epsilon \partial T \cap \varepsilon_{h,\Gamma}$, *we have:*

$$\int_T (\nabla \vec{u} - pI): \nabla \vec{w}_E \, d\Omega = \int_{\partial T} (\nabla \vec{u} - pI)\vec{n}_T. \vec{w}_E \, d\gamma - \int_T (\nabla^2 \vec{u} - \nabla p). \vec{w}_E d\Omega$$

$$= \int_E (\nabla \vec{u} - pI)\vec{n}_T. \vec{w}_E \, d\gamma - \int_T (\nabla^2 \vec{u} - \nabla p). \vec{w}_E d\Omega$$

Since $-\nabla^2 \vec{u} + \nabla p = \vec{f}$ on Ω, and

$$b_0 \vec{u} + \frac{\partial \vec{u}}{\partial n} - \vec{n}p = \vec{t} \quad in \ E \subset \Gamma, \ then$$

$$\int_T (\nabla \vec{u} - pI): \nabla \vec{w}_E \, d\Omega = \int_E (\vec{t} - b_0 \vec{u}) \vec{w}_E \, d\gamma + \int_T \vec{f}. \vec{w}_E \, d\Omega.$$

$$= \int_{\partial T} (\vec{t} - b_0 \vec{u}) \vec{w}_E \, d\gamma + \int_T \vec{f}. \vec{w}_E \, d\Omega.$$

Theorem 5.6. For any mixed finite element approximation (not necessarily inf-sup stable) defined on rectangular grids T_h, the residual estimator η_R satisfies:

$$\|\vec{e}\|_{J,\Omega} + \|\varepsilon\|_{0,\Omega} \leq C_\Omega \eta_R.$$

$$\eta_{R,T} \leq C \left(\sum_{T' \epsilon \omega_T} \{ \|\vec{e}\|_{J,T'}^2 + \|\varepsilon\|_{0,T'}^2 \} \right)^{\frac{1}{2}}.$$

Note that the constant C in the local lower bound is independent of the domain, and

$$\|\vec{e}\|_{J,T'}^2 = |\vec{e}|_{1,T'}^2 + \|\vec{e}\|_{0,\partial T'}^2.$$

Proof. We include this for completeness. To establish the upper bound we let

$[\vec{v}, q] \in H^1(\Omega) \times L^2(\Omega)$ and $\vec{v}_h \in X_h$ be the clement interpolant of $\vec{v}$, then

$$B([\vec{e}, \varepsilon]; [\vec{v}, q]) = B([\vec{e}, \varepsilon]; [\vec{v} - \vec{v}_h, q])$$

$$= \int_\Omega \vec{f}. (\vec{v} - \vec{v}_h) d\Omega - \int_\Omega \nabla \vec{u}_h: \nabla (\vec{v} - \vec{v}_h) d\Omega - \int_\Gamma b_0 \vec{u}_h. (\vec{v} - \vec{v}_h) d\gamma + \int_\Gamma \vec{t}. (\vec{v} - \vec{v}_h) d\gamma + \int_\Omega p_h \nabla. (\vec{v} - \vec{v}_h) d\Omega + \int_\Omega q \nabla. \vec{u}_h d\Omega$$

$$= \sum_{T \epsilon T_h} \{ \int_T \vec{f}. (\vec{v} - \vec{v}_h) d\Omega + \int_T (\nabla^2 \vec{u}_h - \nabla p_h). (\vec{v} - \vec{v}_h) d\Omega - \sum_{E \epsilon \partial T} \int_\Omega \vec{R}_E (\vec{v} - \vec{v}_h) d\gamma + \int_T q. \vec{u}_h d\Omega \}.$$

Thus,

$$|B([\vec{e}, \varepsilon]; [\vec{v}, q])| \leq \sum_{T \epsilon T_h} \{ \|\vec{f} + \nabla^2 \vec{u}_h - \nabla p_h\|_{0,T} \|\vec{v} - \vec{v}_h\|_{0,T} + \sum_{E \epsilon \partial T} \|\vec{R}_E\|_{0,E} \|\vec{v} - \vec{v}_h\|_{0,E} + \|q\|_{0,T} \|\nabla. \vec{u}_h\|_{0,T} \}$$

Then

$$|B([\vec{e}, \varepsilon]; [\vec{v}, q])| \leq \{ (\sum_{T \epsilon T_h} h_T^2 \|\vec{f} + \nabla^2 \vec{u}_h - \nabla p_h\|_{0,T}^2)^{\frac{1}{2}} (\sum_{T \epsilon T_h} \frac{1}{h_T^2} \|\vec{v} - \vec{v}_h\|_{0,T}^2)^{\frac{1}{2}} +$$

$$(\sum_{T \epsilon T_h} \sum_{E \epsilon \partial T} h_E \|\vec{R}_E\|_{0,T}^2)^{\frac{1}{2}} (\sum_{T \epsilon T_h} \sum_{E \epsilon \partial T} \frac{1}{h_E} \|\vec{v} - \vec{v}_h\|_{0,E}^2)^{\frac{1}{2}} +$$

$$(\sum_{T \epsilon T_h} \|q\|_{0,T}^2)^{\frac{1}{2}} (\sum_{T \epsilon T_h} \|\nabla. \vec{u}_h\|_{0,T}^2)^{\frac{1}{2}} \}$$

Using lemma 5.3 then gives:

$$|B([\vec{e}, \varepsilon]; [\vec{v}, q])| \leq C (\sum_{T \epsilon T_h} \{ \|\vec{v}\|_{J,T}^2 + \|q\|_{0,T}^2 \})^{\frac{1}{2}} \times$$

$$\left(\sum_{T \epsilon T_h} \left\{ h_T^2 \|\vec{R}_T\|_{0,T}^2 + \sum_{E \epsilon \partial T} h_E \|\vec{R}_E\|_{0,E}^2 + \|R_T\|_{0,T}^2 \right\} \right)^{\frac{1}{2}}$$

Finally, using the proposition 3.7 gives:

$$\|\vec{e}\|_{J,\Omega} + \|\varepsilon\|_{0,\Omega} \leq C_\Omega \left(\sum_{T \epsilon T_h} \left\{ h_T^2 \|\vec{R}_T\|_{0,T}^2 + \sum_{E \epsilon \partial T} h_E \|\vec{R}_E\|_{0,E}^2 + \|R_T\|_{0,T}^2 \right\} \right)^{\frac{1}{2}}$$

This establishes the upper bound.

Turning to the local lower bound. First, for the element residual part, we have:

$$\int_T \vec{R}_T. \vec{w}_T d\Omega = \int_T (\vec{f} + \nabla^2 \vec{u}_h - \nabla p_h). \vec{w}_T d\Omega$$

$$= \int_\Omega \vec{f}. \vec{w}_T d\Omega - \int_\Omega (\nabla \vec{u}_h - p_h I): \nabla \vec{w}_T d\Omega +$$

$$\int_{\partial T} (\nabla \vec{u}_h - p_h I)\vec{n}. \vec{w}_T d\gamma.$$

Using (41) and (30), gives:

$$\int_T \vec{R}_T. \vec{w}_T d\Omega = \int_T (\nabla \vec{u} - pI): \nabla \vec{w}_T d\Omega - \int_T (\nabla \vec{u}_h - p_h I) \nabla \vec{w}_T d\Omega$$

$$= \int_T (\nabla \vec{e} - \varepsilon I). \nabla \vec{w}_T d\Omega$$

$$\leq (|\vec{e}|_{1,T} + \|\varepsilon\|_{0,T}) |\vec{w}_T|_{1,T}$$

$$\leq C (|\vec{e}|_{1,T}^2 + \|\varepsilon\|_{0,T}^2)^{\frac{1}{2}} h_T^{-1} \|\vec{R}_T\|_{0,T}.$$

In addition, from the inverse inequality (29),

$$\int_T \vec{R}_T. \vec{w}_T d\Omega = \left\| \vec{R}_T. b_T^{\frac{1}{2}} \right\|_{0,T}^2 \geq c \|\vec{R}_T\|_{0,T}^2$$

Thus,

$$h_T^2 \|\vec{R}_T\|_{0,T}^2 \leq C (|\vec{e}|_{1,T}^2 + \|\varepsilon\|_{0,T}^2)$$

$$\leq C (\|\vec{e}\|_{J,T}^2 + \|\varepsilon\|_{0,T}^2). \tag{44}$$

Next comes the divergence part,

$$\|R_T\|_{0,T} = \|\nabla. \vec{u}_h\|_{0,T} = \|\nabla. (\vec{u} - \vec{u}_h)\|_{0,T}$$

$$\leq \sqrt{2} |\vec{u} - \vec{u}_h|_{1,T}$$

$$\leq \sqrt{2} \|\vec{u} - \vec{u}_h\|_{J,T} = \sqrt{2} \|\vec{e}\|_{J,T} \tag{45}$$

Finally, we need to estimate the jump term. For an edge $E \in \partial T \cap \varepsilon_{h,\Omega}$ we have

$$2 \int_E \vec{R}_E. \vec{w}_E d\gamma = \int_{\partial T} (\nabla \vec{u}_h - p_h I)\vec{n}. \vec{w}_E d\gamma$$

$$= \sum_{i=1:2} \int_{\partial T_i} (\nabla \vec{u}_h - p_h I)\vec{n}. \vec{w}_E d\gamma$$

$$= \int_{\omega_E} (\nabla \vec{u}_h - p_h I): \nabla \vec{w}_E d\Omega +$$

$$\sum_{i=1:2} \int_{T_i} (\nabla^2 \vec{u}_h - \nabla p_h). \vec{w}_E d\Omega$$

Using (42) gives:

$$2\int_E \vec{R}_E.\,\vec{w}_E d\gamma = \int_{\omega_E}(\nabla\vec{u}_h - p_h I):\nabla\vec{w}_E d\Omega + \sum_{i=1:2}\int_{T_i}(\nabla^2\vec{u}_h - \nabla p_h).\vec{w}_E d\Omega -$$

$$\int_{\omega_E}(\nabla\vec{u} - pI):\nabla\vec{w}_E d\Omega + \int_{\omega_E}\vec{f}.\vec{w}_E d\Omega$$

$$= \int_{\omega_E}(\nabla\vec{u}_h - p_h I):\nabla\vec{w}_E d\Omega - \int_{\omega_E}(\nabla\vec{u} - pI):\nabla\vec{w}_E d\Omega +$$

$$\sum_{i=1:2}\int_{T_i}\vec{f}.\vec{w}_E d\Omega + \sum_{i=1:2}\int_{T_i}(\nabla^2\vec{u}_h - \nabla p_h).\vec{w}_E d\Omega$$

$$\leq (|\vec{e}|_{1,\omega_E} + \|\varepsilon\|_{0,\omega_E})|\vec{w}_E|_{1,\omega_E} + \sum_{i=1:2}\|\vec{R}_{T_i}\|_{0,T_i}\|\vec{w}_E\|_{0,\omega_E}$$

Using (32) and (33) gives,

$$2\int_E \vec{R}_E.\,\vec{w}_E d\gamma \leq C(|\vec{e}|^2_{1,\omega_E} + \|\varepsilon\|^2_{0,\omega_E})^{\frac{1}{2}}h_E^{-\frac{1}{2}}\|\vec{R}_E\|_{0,E} +$$

$$\sum_{i=1:2}\|\vec{R}_{T_i}\|_{0,T_i}h_E^{\frac{1}{2}}\|\vec{R}_E\|_{0,E}$$

Using (44) gives,

$$2\int_E\vec{R}_E.\,\vec{w}_E d\gamma \leq C(\|\vec{e}\|^2_{J,\omega_E} + \|\varepsilon\|^2_{0,\omega_E})^{\frac{1}{2}}h_E^{-\frac{1}{2}}\|\vec{R}_E\|_{0,E} \qquad (46)$$

Using (31) gives,

$$\int_E\vec{R}_E.\,\vec{w}_E d\gamma = \left\|\vec{R}_E b^{\frac{1}{2}}_E\right\|^2_{0,E} \geq c\|\vec{R}_E\|^2_{0,E}, \text{ and thus}$$

using (46) gives,

$$h_E\|\vec{R}_E\|^2_{0,E} \leq C(\|\vec{e}\|^2_{J,\omega_E} + \|\varepsilon\|^2_{0,\omega_E}). \qquad (47)$$

We also need to show that (47) holds for boundary edges. For an $E \in \partial T \cap \varepsilon_{h,\Gamma}$, we have

$$\int_E\vec{R}_E.\,\vec{w}_E d\gamma = \int_{\partial T}[(b_0\vec{u}_h + (\nabla\vec{u}_h - p_h I)\vec{n} - \vec{t}].\vec{w}_E\, d\gamma$$

$$= \int_{\partial T}(b_0\vec{u}_h - \vec{t}).\vec{w}_E\, d\gamma + \int_{\partial T}((\nabla\vec{u}_h - p_h I)\vec{n}.\vec{w}_E\, d\gamma$$

$$= \int_{\partial T}(b_0\vec{u}_h - \vec{t}).\vec{w}_E\, d\gamma + \int_T((\nabla\vec{u}_h - p_h I):\nabla\vec{w}_E\, d\Omega +$$

$$\int_T(\nabla^2\vec{u}_h - \nabla p_h I).\vec{w}_E\, d\Omega.$$

Using (43) and (31), gives

$$\int_E\vec{R}_E.\,\vec{w}_E d\gamma = \int_{\partial T}(b_0\vec{u}_h - \vec{t}).\vec{w}_E\, d\gamma + \int_T((\nabla\vec{u}_h - p_h I):\nabla\vec{w}_E\, d\Omega + \int_T(\nabla^2\vec{u}_h - \nabla p_h I).\vec{w}_E\, d\Omega$$

$$= -\int_T((\nabla\vec{e} - \varepsilon I):\nabla\vec{w}_E\, d\Omega - \int_{\partial T}b_0\vec{e}.\vec{w}_E\, d\gamma +$$

$$\int_T\vec{R}_T.\vec{w}_E\, d\Omega$$

$$\leq C(\|\vec{e}\|_{J,T} + \|\varepsilon\|_{0,T})\|\vec{w}_E\|_{J,T} +$$

$$\|\vec{R}_T\|_{0,T}h_E^{\frac{1}{2}}\|\vec{R}_E\|_{0,E}.$$

Using (44) and (34), we obtain

$$\int_E\vec{R}_E.\,\vec{w}_E d\gamma \leq C(\|\vec{e}\|^2_{J,T} + \|\varepsilon\|^2_{0,T})^{\frac{1}{2}}h_E^{-\frac{1}{2}}\|\vec{R}_E\|_{0,E}. \qquad (48)$$

Using (31)

$$\int_E\vec{R}_E.\,\vec{w}_E d\gamma = \left\|\vec{R}_E b^{\frac{1}{2}}_E\right\|^2_{0,E} \geq c\|\vec{R}_E\|^2_{0,E},$$

and thus using (48) gives,

$$h_E\left\|\vec{R}_E\right\|^2_{0,E} \leq C(\|\vec{e}\|^2_{J,T} + \|\varepsilon\|^2_{0,T}). \qquad (49)$$

Finally, combining (44), (45), (47) and (49) establishes the local lower bound.

Remark 5.7. Theorem 5.6 also holds for stable (and unstable) mixed approximations defined on a triangular subdivision if we take the obvious interpretation of ω_T. The Proof is identical except for the need to define appropriate element and edge bubble functions.

VI. FIGURES AND TABLES

In this section some numerical results of calculations with mixed finite element Method and ADINA System will be presented. Using our solver, we run the test problem driven cavity flow [14, 16] with a number of different model parameters.

Example. Square domain, enclosed flow boundary condition.
This is a classic test problem used in fluid dynamics, known as driven-cavity flow. It is a model of the flow in a square cavity with the lid moving from left to right. Let the computational model:
$(y = 1; -1\leq x\leq 1$ / $u_x=1-x^2)$, a regularized cavity. With these data, see that the (D+N) condition is satisfied, just take b_0 a real number very large and $\vec{t} = (b_0(1 - x^2); 0)$ on
$\Gamma_1 =(y = 1; -1\leq x\leq 1)$ and $\vec{t} = (0; 0)$ on the other three boundary of the square domain.
The streamlines are computed from the velocity solution by solving the Poisson equation numerically subject to a zero Dirichlet boundary condition.
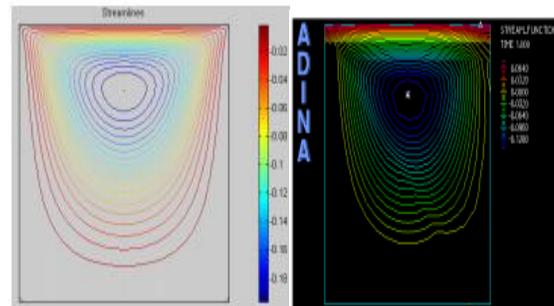
Fig.1. Uniform streamline plot by MFE (left) associated with a 64-64 square grid, Q_1-P_0 approximation, and uniform streamline plot (right) computed with ADINA system.
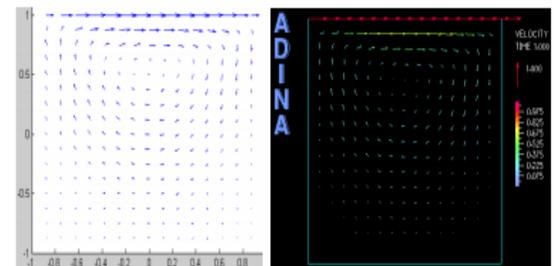
Fig.2. Velocity vectors solution by MFE (left) associated with a 64-64 square grid, Q_1P_0 approximation and velocity vectors solution (right) computed with ADINA system.
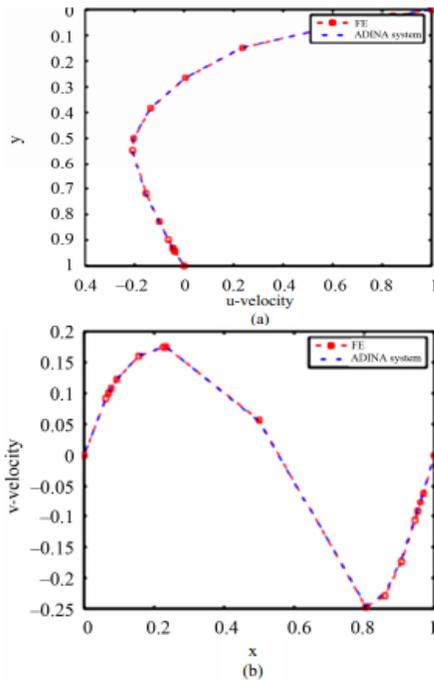
Fig.3. The velocity component u at vertical center line (a), and the velocity component v at horizontal center line (b) with a 129×129 grid.

Figure 3 shows the velocity profiles for lines passing through the geometric center of the cavity.

These features clearly demonstrate the high accuracy achieved by the proposed mixed finite element method for solving the Stokes equations in the lid-driven squared cavity.
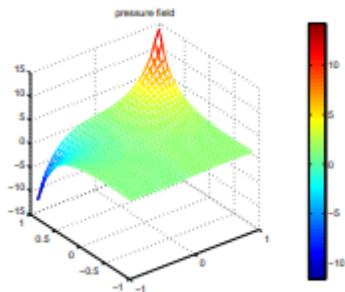
Fig.4. Pressure plot for the flow with a 64-64 square grid.

VII. CONCLUSION

In this work, we were interested in the numerical solution of the partial differential equations by simulating the flow of an incompressible fluid. We introduced the Stokes equations with a boundary condition of type (D+N).

The weak formulation obtained is a problem of saddle point type. We have shown the existence and uniqueness of the solution of this problem. We used the discretization by mixed finite element method with a posteriori error estimation of the computed solutions. For the test of driven-cavity flow, the particles in the body of the fluid move in a circular trajectory.

Our results agree with Adina system.

Numerical results are presented to see the performance of the method, and seem to be interesting by comparing them with other recent results.

REFERENCES

[1] Alexandre Ern, *Aide-mémoire Eléments Finis*, Dunod, Paris, 2005.

[2] P.A. Raviart, J. Thomas, *Introduction l'analyse numérique des équations aux dérivées partielles,* Masson, Paris, 1983.

[3] R. Dautray et J.L. Lions, *Mathematical Analysis and Numerical Methods for science and Technology,* Vol.4. Integral equations and numerical methods, Springer-Verlag, Berlin, Allemagne, 1990.

[4] F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Method,* Computational Mathematics, Springer Verlag, New York, 1991.

[5] P.A. Raviart, J.M. Thomas, A mixed finite element method for second order elliptic problems, *in: Mathematical Aspects of Finite Element Method.* Lecture Notes in Mathematics, Springer, New York, 1977, 292-315.

[6] G. Chavent and J. Jaffré, *Mathematical Models and Finite Elements for Reservoir Simulation,* Elsevier Science Publishers B.V, Netherlands, 1986.

[7] J. Roberts and J.M. Thomas, Mixed and Hybrid methods, *Handbook of numerical analysis II, Finite element methods 1, P. Ciarlet and J. Lions,* Amsterdam, 1989.

[8] R.E. Bank, A. Weiser. Some a posteriori error estimators for elliptic partial differential equations, *Mathematics of Computation, 44(170),* 1985, 283-301.

[9] M. Ainsworth, J. Oden, A posteriori error estimates for Stokes' and Oseen's equations, *SIAM Journal of Numerical Analytic, 34(1),* 1997, 228-245.

[10] R.E. Bank, B. Welfert, A posteriori error estimates for the Stokes problem, *SIAM Journal of Numerical Analytic, 28(3),* 1991, 591-623.

[11] C. Carstensen, S.A. Funken, A posteriori error control in low-order finite element discretizations of incompressible stationary flow problems, *Mathematics of Computation, 70(236),* 2001, 1353-1381.

[12] D. Kay, D. Silvester, A posteriori error estimation for stabilized mixed approximations of the Stokes

equations, *SIAM J. Sci. Comput. 21(4),* 1999, 1321-1336.

[13] R. Verfurth, A posteriori error estimators for the Stokes equations, *Numer. Journal of Numerical Mathematics, 55(3),* 1989, 309-325.

[14] H. Elman, D. Silvester, A. Wathen, *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics,* Oxford University Press, Oxford, 2005.

[15] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques,* Wiley-Teubner, Chichester, 1996.

[16] D.H. Wu, I.G. Currie, Analysis of a posteriori error indicator in viscous flows, *International Journal of Numerical Methods for Heat and Fluid Flow, 12(228),* 2002, 1347–1378.

[17] Y. He, A. Wang, L. Mei, Stabilized finite-element method for the stationary Navier-Stokes equations, *Journal of Engineering Mathematics, 51(4),* 2005, 367-380.

[18] V. John, Residual a posteriori error estimates for two-level finite element methods for the Navier-Stokes equations, *Journal of Numerical Mathematics, 37(4),* 2001, 503-518.

[19] E. Creuse, G. Kunert, S. Nicaise, A posteriori error estimation for the Stokes problem: Anisotropic and isotropic discretizations, *Mathematical Models and Methods in Applied Sciences, 14(9),* 2004, 1297-1341.

[20] P. Clement, Approximation by finite element functions using local regularization, *RAIRO. Anayse, Numérique, 2(9),* 1975, 77-84.

[21] V. Girault and P.A. Raviart, *Finite Element Approximation of the Navier-Stokes Equations,* Springer-Verlag, Berlin Heiderlberg New York, 1981.

Design and Implementation Of Low Power CMOS Radio Receiver

[1]Neha Agarwal , [2]Dwijendra Parashar

[1](Department of Electronics and Communication, Laxmi devi Institute of Engg. & Technology Alwar)India
[2](Department of Electronics and Communication, Shobhit University Meerut, India

Abstract
A single-chip CMOS Global Positioning System (GPS) radio has been integrated using only a couple of external passive components for the input matching network and one external reference for the synthesizer. This paper explores architectural and design techniques for CMOS wireless receivers through the vehicle of the GPS system. This system comprises 24 satellites in low earth orbit that continuously broadcast their position and local time. Through satellite range measurements, a receiver can determine its absolute position and time anywhere on Earth, as long as four satellites are within view. Examples of such applications include automotive or maritime navigation, intelligent hand off algorithms in cellular telephony, and cellular emergency services, to name a few. To enable a cheap, low-power CMOS GPS solution, this work develops a receiver architecture that lends itself to complete integration. To implement this architecture, two major foci are the design of low-noise amplifiers (LNAs) and power efficient active filters in CMOS technologies. The realization of a 2.4dB noise figure, differential LNA is with only 12mW power consumption in a 0.5µm CMOS technology. Another focus is on the power efficient implementation of wide dynamic range active filters. In such filters, the design of the transconductor element is critical, and techniques for evaluating transconductor architectures are presented. An application of these ideas are to the GPS receiver problem results in a 10mW, 60dB peak spurious free dynamic range (SFDR) active filter with 3.5MHz bandwidth. These advances enable the realization of an 115mW CMOS GPS receiver that includes the complete RF and analog signal path, frequency synthesizer and A/D converters.

Keywords-   *Introduction, Radio receiver, GPS system architecture, Chip design, CMOS Mixer, Conclusion, Reference.*

I.   Introduction

GLOBAL Positioning System (GPS) receivers for the consumer market require solutions that are compact, cheap, and low power. Manufacturers of cellular telephones, portable computers, watches, and other mobile devices are looking for ways to embed GPS into their products. Thus, there is a strong motivation to provide highly integrated solutions at the lowest possible power Consumption. GPS radios consist of a front-end and a digital baseband section incorporating a digital processor. While for the baseband processor, cost-reduction reasons dictate the use of the most dense digital CMOS technology, for the front-end, the best option in terms of power consumption is a SiGe BiCMOS technology. This explains why several commercial GPS radios consist of dual or multichip systems using the best technology option for the front-end and baseband processor. On the other hand, the implementation of a stand-alone GPS radio into a single chip in CMOS technology is appealing in terms of cost, and would speed up the integration of GPS capabilities into mobile products. This motivated the development of GPS macro blocks and radios in CMOS technology. However, the cost effectiveness of this solution depends on both reduction of external components and die area of the GPS radio. Since the silicon area of RF CMOS circuits, including on-chip inductors, does not shrink at the same rate as technology scaling, the reduction of the total cost poses a severe challenge. Typically, GPS radios are implemented in bipolar or BiCMOS processes and cannot be integrated with the digital signal processor chip due to higher cost. Furthermore, they use multi-down conversion receiver architectures that require off-chip filters, adding to the footprint and cost.

The Global Positioning System (GPS) is a satellite-based navigation system made up of a network of 24 satellites placed into orbit by the U.S. Department of Defense. GPS was originally intended for military applications, but in the 1980s, the government made the system available for civilian use. GPS works in any weather conditions, anywhere in the world, 24 hours a day. There are no subscription fees or setup charges to use GPS. All satellites transmit at both frequencies and their signals are distinguished by different Gold codes used to spread the signal, where most commercial GPS receivers use the L1 signal only. The system is based on time-of-arrival (TOA) of signals from the visible satellites. Using the location and time information sent by satellites and TOA from at least four visible satellites, four equations can be solved for altitude, latitude, longitude, and time.

II.   Radio Receiver Architecture

The advent of wireless communications at the turn of the 20th century marked the beginning of a technological era in which the nature of communications would be radically altered. The ability to transmit messages through the air would soon usher in radio and television broadcasting and wireless techniques would later and application in many of the mundane tasks of everyday life. Today, the widespread use of wireless technology conveys many benefits that are easily taken for granted. From cellular phones to walkie-talkies; from broadcast television to garage door openers; from aircraft radar to hand-held GPS navigation systems, radio technology pervades modern life. At the forefront of emerging radio applications lies modern research on the integrated radio receiver. The goal of miniaturization made possible by integrated circuit technologies holds the promise of portable, cheap and robust radio systems, as exampled by the advent of cellular telephony in the mid-1980's. As miniaturization continues, embedded radio applications become possible where the features of multiple wireless systems can be brought to bear on a particular problem. One example is the use of a GPS receiver in a cellular telephone to permit the expedient dispatch of emergency service personnel to the caller's exact location.
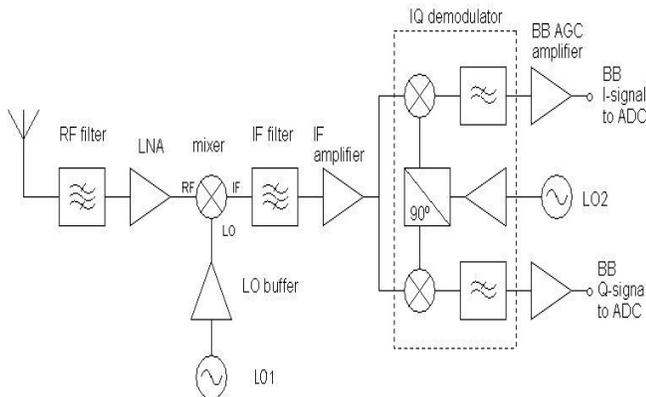
Fig-1 Super Heterodyne Receiver.

Figure1 makes use of a parallel data detector concept. There are two types of detector such serial data and parallel data. Both detector concepts are generally applicable for demodulation. The main difference is that the serial detector makes use of real signal representation whereas the parallel detector uses complex signal representation. As a result the

Serial detector is generally less complicated and lower power consumption is potential. However, the parallel detector generally provides for better performance in fading environments than does the serial detector and as a result the parallel detector is the most widely used. Also, the use of the parallel data detector concept of Figure1 generally improves on the image rejection. Ideally, the complex signal representation provided by the parallel detector allows for

Complete separation of the desired signal and the corresponding image signal. By prober combining of the I and Q signals it is possible to relocate the desired and the image signal to positive and negative frequencies, respectively. This, however, requires exact matching of the I and Q signal branches which is not possible in a practical implementation. One way of illustrating the effect of I/Q mismatch is to consider a complex input signal consisting of unmodulated single tone signals,

$$s(t) = \cos(2\pi f_{IM} t) + j\alpha \sin(2\pi f_{IM} t + \theta)$$

$$s(t) = \tfrac{1}{2}\,[e^{j2\pi f l M t}(1 + \alpha.e^{j\theta})] + \tfrac{1}{2}[e^{-j2\pi f l M t}(1 - \alpha e^{j\theta})] \quad (1)$$

As the amplitude α and phase θ errors are relative values these can be assigned to just one of the signal. The term $e^{j2\pi f l M t}$ and $e^{-j2\pi f l M t}$ can be view as the desired and image signal respectively. Hence, from Equation (1) it is seen that in case of ideal matching, i.e. $(\alpha, \theta) = (1; 0)$, the image signal component is cancelled. By comparison of the amplitude terms of the desired and the image signal components he approximate amount of I/Q image discrimination.

A.   Radio Spectrum

The goal of any radio receiver is to extract and detect selectively a desired signal from the electromagnetic spectrum. This selectivity" in the presence of a plethora of interfering signals and noise is the fundamental attribute that drives many of the tradeoffs inherent in radio design. Radio receivers must often be able to detect signal powers as small as a femtowatt while rejecting a multitude of other signals that may be twelve orders of magnitude larger! Because the electromagnetic spectrum is a scarce resource, interfering signals often lie very close to the desired one in frequency, thereby exacerbating the task of rejecting the unwanted signals.

B.   Classical Receiver Architecture

The design of wireless receivers is a complex, multi-faceted subject that has a fascinating history. In this section, we will explore many of the fundamental issues that arise in receiver design through the vehicle of historical examples. These early receiver architectures illustrate an increasing level of sophistication in response to the need for improved selectivity at ever-greater frequencies.

C.   Crystal detector

The received signal from the antenna is band pass filtered and immediately rectified by a simple diode. If a sufficiently strong amplitude modulated radio signal is received, the rectified signal will possess an audio frequency component that can be heard directly on a pair of high-impedance headphones. The desired radio channel can be selected via a variable capacitor (or condenser, according to the terminology of the day). Remarkably, this

radio does not require a battery; the received signal energy drives the headphones directly without amplification.

III.   GPS System Architecture

SUCCESSFUL radio designs begin with good architectural choices. Unfortunately, there is no radio architecture panacea. Rather, it is essential to select the approach best suited for the task at hand. In this section, we turn our attention to selecting the GPS radio architecture that will permit the maximum level of integration while minimizing power consumption. We begin with the details of the GPS system itself. As will be shown, the GPS system possesses certain unique features that make it particularly well suited for integration.

To motivate the architectural choices described in this chapter, it is important to consider some details of the received GPS signal spectrum. The GPS system uses a direct-sequence spread spectrum technique for broadcasting navigation signals. In such an approach, the navigation data signal is multiplied by a pseudo-random bit sequence (PRBS) code that runs at a much higher rate than the navigation symbol rate. This higher rate is commonly referred to as the "chip rate" of the code. The PRBS codes used in the GPS system are Gold codes that have two possible values (±1) at any given time. Thus, when a code is multiplied by itself, the result is a constant value; however, when two different codes multiply each other, the result is another PRBS sequence. This property can be used to separate overlapping received signals from multiple satellites into distinct data paths for navigation processing. In principle, by multiplying the received signal by a particular satellite's PRBS code, the receiver can recover data from that satellite alone while signals from other satellites pass through with the appearance of pseudo-random noise. Hence, with a unique PRBS code assigned to each satellite, all satellites can broadcast at the same frequency without substantially interfering with each other.

The GPS satellites broadcast navigation signals in two bands: the L1 band, which is centred at 1.57542GHz, and the L2 band, centred at 1.2276GHz. Each satellite broadcasts two different direct-sequence spread-spectrum signals. These are known as the P code (or precision code) and the C/A code (or coarse acquisition code). The P code is broadcast in both frequency bands, while the C/A code is broadcast only in the L1 band. Note that the centre frequencies of the L1 and L2 bands are both integer multiples of 10.23MHz, which is the chip rate of the P code signal. In contrast, the C/A code uses a lower chip rate of 1.023MHz. The P code is intended for military use and is much more difficult to detect, in part because it uses a spreading code that only repeats at 1-week intervals. In addition, the P code is encrypted to restrict its use to authorized (military) users. For this reason, the C/A code is of primary interest in commercial applications.

The spectrum of the GPS L1 band shows that the C/A code and the P code occupy the same 20-MHz spectrum allocation, but their main lobes have different bandwidths due to the different code chip rates. In particular, the C/A code has a main lobe width of 2MHz while the P code has a width of 20MHz. The outlying lobes of the P code are truncated by appropriate filtering so that the entire GPS broadcast fits neatly within the 20-MHz allocation. The immunity to interference that is gained when using the spread-spectrum technique is related to the ratio of the chip rate to the symbol rate.
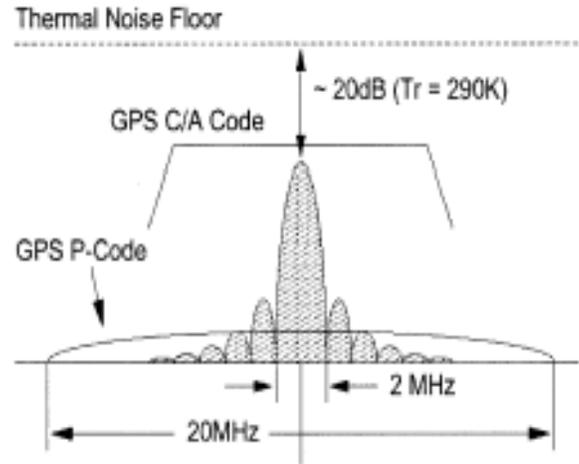
Fig-2 GPS L1 band signal spectrum.

This ratio, called the processing gain, gives an indication of the improvement in SNR that occurs when a signal is de-spread". For the GPS C/A code, the symbol rate is a mere 50Hz. Thus, the processing gain is given by

$$G_p = 10\log\left(\frac{f_c}{f_b}\right) = 43\text{dB} \tag{2}$$

Where f_b the symbol rate of the C/A code and f_c is the chip rate. The received signal power is typically 130dBm at the antenna of a GPS receiver. If we assume that we are primarily interested in the 2-MHz main lobe of the C/A code, the noise power in this 2-MHz bandwidth is simply given by kTB≈-111dBm (T = 290K). Hence, the received SNR at the antenna is about 19dB. Once the signal from a given satellite is correlated with its PRBS code, the bandwidth is reduced to only 100Hz. Thus, the post correlation SNR improves by the processing gain of the system. So, with an antenna temperature of 290K and an otherwise noiseless receiver, the post-correlation SNR would be about 24dB.

IV.   Chip Design

As stated, the overall design has been geared to a high level of integration and reduction of silicon area at the lowest possible power consumption. Below, the detailed design choices in the various sections are described.

a)   RF Section-

The LNA has been designed to have a very low noise since it sets a lower bound for the total receiver sensitivity. A high voltage gain is necessary to sufficiently reduce the noise contribution of the following mixers. A common source configuration with inductive degeneration provides high voltage gain and low NF, as shown in Fig. 3. In fact, in a narrow band, this structure allows achieving a noise factor close to the theoretical minimum.
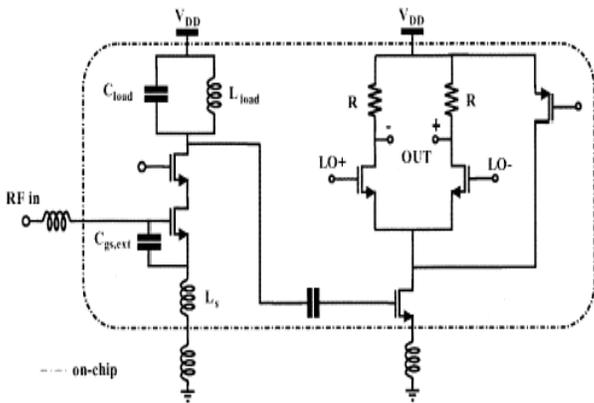
Fig-3 Low noise amplifier and mixer

A single-ended LNA has been preferred to a balanced one to reduce power consumption and silicon area. The input of the mixer is still single ended, but from its output, the signal is taken in a differential fashion. In this topology, at a given frequency, there is an optimum device size for which the sum of gate induced and thermal noise has a minimum. Because of the strong sensitivity of the gate-induced current noise to the intrinsic gate capacitance (it follows a square law), an improvement can be obtained with the introduction of an additional capacitance C_{gset} placed in parallel to the intrinsic gate capacitance C_{gs} of the input transistor. The insertion of this capacitance adds a degree of freedom to play with to achieve a better compromise between thermal and induced-gate noise. Therefore, a new optimum condition, with a lower noise figure minimum, can be achieved. This is paid by a slightly lower transconductance gain.

The LNA is followed by the - mixers that are ac-coupled to the LNA and are based on a modified Gilbert cell. The mixers can be directly driven by the on-chip frequency synthesizer or by a single external local oscillator (LO) signal that drives an integrated RF poly phase filter. Improved linearity and reduced noise are achieved by subtracting dc current from the switching pair. The load is a simple resistor. The current consumption is 1.5 mA for each mixer with an input P1 dB of 12 dBm.

b)   IF Section-

After down conversion, the signal is amplified using a variable-gain amplifier (VGA) with 20-dB gain programmability. A second-order integrated passive poly phase filter has been used to recombine the I and Q signal path. The poly phase filter is an *RC* structure with inputs and outputs symmetrically disposed. The relatively small ratio between the signal band and IF frequency allows building the combiner as the cascade of two *RC* passive poly-phase filters. A rejection of 30 dB across the 2-MHz band is achieved for ±20% *RC* time constant spread.

The IF filter is centered at 9.45 MHz To fit the 2-MHz GPS band, even in presence of component values variations, the nominal transfer function features a larger bandwidth (6 MHz) than the one needed (2 MHz). However, a ripple in the GPS band (8.45–10.45 MHz) lower than 0.5 dB is guaranteed in any case. To optimize the power consumption for a given linearity and noise, an active *RC* solution has been chosen. The filter is built as a cascade of a band pass and a low-pass cell, implementing a fourth-order transfer function. The filter also provides an antialiasing function before the baseband ADC, assuring 20-dB attenuation at 28 MHz
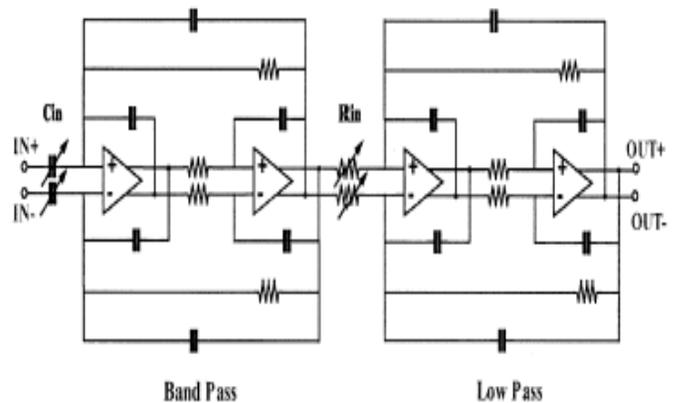
Fig-4 IF Filter

c)   Synthesizer

The synthesizer, depicted in Fig, provides LO quadrature signals for the image-reject mixer and two clock signals needed to synchronize the correlator inside the external baseband processor. As for the previously described blocks, the main concern has been a high level of integration and reduction of silicon area at the lowest possible power consumption. The requirement of high integration and the need to reduce risks of LO pulling due

to off-chip components have driven the choice of a fully integrated voltage-controlled oscillator (VCO) and loop filter. For this application, we took advantage of the low requirements on phase noise and spurious rejection using a ring oscillator within a wide-band phase-locked loop (PLL) instead of an inductance-capacitance (*LC*) VCO. This choice resulted in a dramatic reduction of the silicon area, since it does not require integrated inductors or varactors. Furthermore, the ring oscillator directly provides *I-Q* quadrature LO signals needed by the image-reject mixer and simplifies the portability of the GPS radio to a pure CMOS process with associated low-quality factor on-chip inductors. The use of a ring oscillator requires a wide-band PLL to reduce its contribution to the total phase noise. The added benefit of this choice is the integration of the loop filter in a limited silicon area. This is highly desirable to reduce the risk of VCO pulling due to external interferences coupled through the wire bonding. A 500-kHz band is implemented with 12 pF in parallel with the series of 300 pF and 6.4 kΩ.

The synthesizer is comprised of a crystal oscillator that generates the 18.414-MHz reference frequency using an off-chip crystal. Its output is divided by four and used as the comparison frequency. The VCO generates the quadrature local oscillator (LO) signals for the receiver mixers. The LO signal is divided by 342 and input to the PFD along with the comparison frequency.

The PFD has a delay element in its feedback loop to prevent a dead- zone region in the charge pump. The resulting filtered charge-pump output controls the VCO. Here, an off-chip second-order filter was used. GPS is a single-frequency system and a start-up time of less than 5 ms is acceptable. Therefore, synthesizer-settling time was not an important design factor.
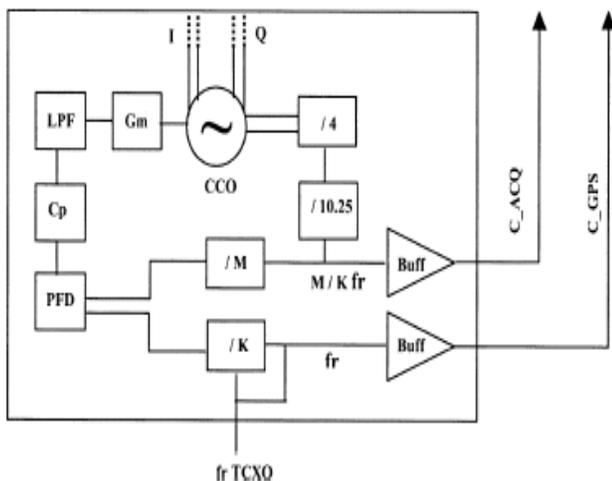
Fig-5 PLL Synthesizer

V.   CMOS Mixer

Mixer mixes or multiplies two signals to get a resultant output signal. A time domain multiplication is a convolution in frequency domain so, in the transceiver, mixers perform frequency translation or frequency up-conversion and down-conversion by multiplying an RF input signal, with frequency wRF, and an LO signal, with frequency wLO, present at mixers the RF and the LO ports, respectively. the local oscillator is coupled to the center taps of the two transformers to control which pair of diodes is forward biased. The different mode of operation for the diode ring mixer in which the LO drive is transformer-coupled to the ring. In this case, a different pair of diodes is activated on each phase of the LO that couples the other transformer to the center taps of the two secondary. This connection is widely used when the RF and LO signals are at high frequencies and the IF output is at a relatively low frequency. Center tapping the IF port permits the use of smaller self inductances in the transformer secondary, which only need to present large impedances at the LO and RF frequencies. While diode ring mixers operate on a voltage switching principle, the Gilbert mixer operates on a current switching principle. The mixer employs two bipolar current switches whose outputs are connected in opposition. On one phase of the LO, the input current owes to the output through the outer pair of devices, while on the opposite phase, the input current is diverted to the opposite outputs via the inner pair of devices. If the bases of the transistors are driven symmetrically, the emitters lie at points of symmetry so that, in principle, no local oscillator signal couples to the input or output ports. So, just as in the diode ring case, the Gilbert mixer is nominally a double-balanced mixer.

VI.   Conclusion

The vast majority of integrated GPS receivers use a standard super heterodyne architecture with a number of off-chip components, particularly passive IF filters. It was demonstrated how the detailed nature of the GPS signal spectrum presents an opportunity for a low-IF receiver architecture that offers the benefit of complete integration of the receiver signal path. Because the I/Q matching requirements are relaxed in this architecture, adequate image rejection is readily obtained without trimming or calibration. With the goal of minimum power consumption in mind, Beginning with the low-noise amplifier, Low noise operation in CMOS by including an oft-neglected noise source: induced gate noise. This power-constrained" optimization demonstrates that excellent noise performance can be achieved in CMOS with small numbers of milliwatts while delivering a good input impedance match to the off-chip 50 world. This theoretical development enables a 2.4dB noise figure for a differential LNA with only 4.9mA of bias current.
The double-balanced CMOS voltage mixer consumes exceptionally little power with no bias current required in the mixer core. Although conversion gain is a concern in

this architecture, a careful analysis demonstrates that higher conversion gains can be obtained by reactively terminating the IF port of the mixer. A noise figure analysis further demonstrated that SSB noise figures on the order of 6dB are easy to obtain. Finally, the linearity of the mixer is primarily limited by the magnitude of the LO drive. Thus, the CMOS voltage mixer achieves wide dynamic range and excellent noise figure with no static power consumption In the future, process scaling may necessitate a reduction in supply voltage for integrated radio receivers. This trend will present a number of challenges for wide dynamic range receiver design.

Acknowledgement:

VII.   REFERENCES

- Aldert van der Ziel, Noise in Solid State Devices and Circuits, John Wiley & Sons, New York, 1986.

- A. Abidi, \High-frequency noise measurements on FET's with small dimensions,"IEEE Transactions on Electron Devices, vol. ED-33, no. 11, pp. 1801-1805, Nov. 1986.

- Norman G. Einspruch, Ed., VLSI Electronics: Microstructure Science, vol. 18, pp. 1{37, Academic Press, New York, 1989.

- Bradford W. Parkinson, \Introduction and heritage of NAVSTAR, the global positioning system," In Parkinson and Spilker [104], pp. 3-28.

- Guglielmo Marconi, \Radio telegraphy," Proceedings of the IRE, vol. 10, pp. 215-238, Aug. 1922.

- Thomas H. Lee, The Design of CMOS Radio Frequency Integrated Circuits, Cambridge University Press, 1998.

- R. H. Marriott, United States radio development," Proceedings of the IRE, vol. 5, pp. 179-198, June 1917.

- Guglielmo Marconi, \Radio communication," Proceedings of the IRE, vol. 16, pp. 40-69, Jan. 1928.

- Michael Riordan and Lillian Hoddeson, Crystal Fire: The Birth of the Information Age, W. W. Norton & Company, 1997.

- Lee de Forest, \The audion {detector and amplifier," Proceedings of the IRE, vol. 2, pp. 15{36, Mar. 1914.

- Haraden Pratt, \Long range reception with combined crystal detector and audion amplifier," Proceedings of the IRE, vol. 3, pp. 173{183, June 1915.

- Edwin H. Armstrong, \Operating features of the audion," Electrical World, no. 24, pp. 1149{1152, Dec. 1914.

- John L. Hogan, Jr., \The heterodyne receiving system and notes on the recent Arlington-Salem tests," Proceedings of the IRE, vol. 1, pp. 75{97, July 1913.

- Benjamin Liebowitz, \The theory of heterodyne receivers," Proceedings of the IRE, vol. 3, pp. 185{204, Sept. 1915.

- Edwin H. Armstrong, \Some recent developments in the audion receiver," Proceedings of the IRE, vol. 3, pp. 215{247, Sept. 1915.

An Improved Collude Attack Prevention for Data Leakage

[1]Keerthana.P, [2]Narmadha.R.P

[1](Final ME (CSE), Sri Shakthi Institute Of Engineering and Technology, Coimbatore) India.
[2](Assistant Professor, Sri Shakthi Institute Of Engineering and Technology, Coimbatore) India

ABSTRACT:

A data distributer needs to secure the sensitive data at the time of distributing to the different agents. The sensitive data should be transferred in a secured way and it should not be leaked to some other persons. The existing system uses the technique called Watermarking to prevent the leakage. In the other techniques, the fake objects are attached with the real objects to detect the leakage and the guilty agent who leaking the data. The disadvantages of this system are the originality of the data is lost and some data cannot be transformed. In the proposed system, it concentrates on preventing two agents to compare and extract the fake objects. Symmetric Inference Model (SIM) is introduced for cases where agents can collude and identify fake tuples. In this technique it uses the symmetric inference graph approach for the symmetric inference model that represents the possible colluding attacks from any agents to the different data allocation strategies. SIM represents dependent and semantic relationships among attributes of all the entities in the information system. This prevents the agents from comparing their data with one another to identify fake objects.

Index terms- Allocation strategies, data leakage, data privacy, fake records, leakage model, Symmetric inference model.

I.INTRODUCTION

While doing business, sometimes we have to transfer the sensitive data to trusted third parties. For example, a company may have partnerships with many companies. They need to share their sensitive data. We call the customer as agents and the owner of data as distributor. Our main aim is to detect when the data is leaked by agents and also to find who had leaked out the data.

In some cases the original sensitive data is modified and handed over to the agents. The technique used to modify the data and make "less sensitive" is perturbation. But there are some cases like medical researches and payroll in which the ranges of data should not be modified as they need accurate data for treating patients in case of medical and correct account number for salary calculation in payroll case.

At first the leakage is controlled by the method called watermarking. In that a unique code will be embedded with the distributed copy. If the data is found in unauthorized parties we can identify the leaker. The disadvantage in this is the original copy need to be modified. And also in some cases the watermarks can be destroyed if the recipient is malicious.

After giving certain set of objects to agents, the distributor may find some data in unauthorized place. For example if it s found in some web site, we can identify the leaker with the help of cookies. But with a single cookie we can't proof his leakage. So if we have strong information like four or five cookies we can decide what to do with that agent.

In this paper we develop a model to protect the data if the agents decide to meet and find the fake objects by comparing with their records. Here the algorithms are designed to distribute the data. Then the fake objects are attached with the original data and distributed. The fake objects will look like the original to the agents. They will act as watermark to find the leaker of data. It will turn up and intimate the distributor, and then the distributor can be confident that the agent is guilty agent.

Entities and Agents

The distributor wants to share some of the objects with a set of agents U1, U2, . . . , Un, but does not wish the objects to be leaked to other third parties. The objects in T could be of any type and size, e.g., they could be tuples in a relation, or relations in a database.

An agent Ui receives a subset of objects Ri in T, determined either by a sample request or an explicit request:

- Sample request Ri = SAMPLE(T,mi): Any subset of mi records from T can be given to Ui.
- Explicit request Ri = EXPLICIT(T,condi): Agent Ui receives all T objects that satisfy condition.

Fake Objects

The distributor may be able to add fake objects to the distributed data in order to improve his effectiveness in detecting guilty agents. Although we do not deal with the implementation of CREATE FAKE OBJECT(), we note that there are two main design options. The function can either produce a fake object on demand every time it is called or it can return an appropriate object from a pool of objects created in advance.

In section 2 we start by describing the relative work whereas in section 3 discuss briefly about the analysis of agent guilt model and the various data allocation strategies respectively. The experimental results and the technique to calculate the probability of agent colluding attacks is explained in section 4.

II. AN OVERVIEW OF RELATED WORK

The guilt detection approach we present is related to the data provenance problem: tracing the lineage of S objects implies essentially the detection of the guilty agents. Tutorial provides a good overview on the research conducted in this field. Suggested solutions are domain specific, such as lineage tracing for data warehouses and assume some prior knowledge on the way a data view is created out of data sources.

In the watermarking relational databases paper, the watermark is inserted in the range of bits. The bits get replaced to act as the watermark. It helps to prevent the attack from the malicious attack and bit flipping attack and so on. The disadvantage is the watermarking is not developed for the non-numeric attributes.

The goal of watermarking is to insert the mark in the object without destroy the value of the object and it is difficult for the adversary to remove or alter the mark beyond detection without destroying the original value. Database semantics and structured data are the challenges faced during this process.

The protection is also achieved through the process of k-anonymity. In this k-anonymity provides privacy protection by guaranteeing that each record relates to at least k individuals even if the released records are directly linked to external information. It provides a formal presentation of achieving k-anonymity using generalization and suppression. Generalization involves replacing a value with a less specific but semantically consistent value. Suppression involves not releasing a value at all.

Finally, there are also lots of other works on mechanisms that allow only authorized users to access sensitive data through access control policies.

III. AGENT GUILT MODEL ANALYSIS

For finding the guilty agent, the probability is to be calculated. For example if we are giving data to two agents then the probability of the agent to be guilty is 0.5. To estimate how likely it is that a system will be operational throughout a given period, we need the probabilities that individual components will or will not fail. A component failure in our case is the event that the target guesses an object of S. The component failure is used to compute the overall system reliability, while we use the probability of guessing to identify agents that have leaked information. The component failure probabilities are estimated based on experiments. Similarly, the component probabilities are usually conservative estimates, rather than exact numbers.

Data Allocation Strategies

The main problem is to allocate the data to the agents with the high probability of finding the guilty agent. For that four possibilities are found. As illustrated in Fig. 1, there are four instances of this problem we address, depending on the type of data requests made by agents whether it is explicit data request or sample data request and whether the "fake objects" are allowed or not.
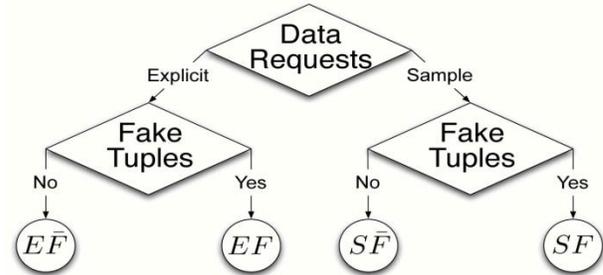
Fig. 1. Leakage problem instances.

While allocating data to the agents, the constraint is to satisfy the agent's request by providing the entire request that are available and the objective is to detect the agent if he leaks the data.

A. Explicit data request
Explicit Data Request with e-random

In this model we present an approach of explicit data request based on e-random. Here we combine the allocation of the explicit data request with the agent selection of e-random. We use e-random as our baseline in our comparisons with other algorithms for explicit data requests. Initially we the finds agents that are eligible to receiving fake objects in O(n) time. Then, the algorithm creates one fake object in every iteration and allocates it to random agent. The main loop takes O(B) time. Hence, the running time of the algorithm is O(n + B).

Algorithm 1. Allocation for Explicit Data Requests (EF)

Input: $R_1, \ldots R_n$, $cond_1, \ldots, cond_n$, $b_1, \ldots, b_n$, B
Output: $R_1, \ldots, R_n$, $F_1, \ldots, F_n$
1: R ← φ . Agents that can receive fake objects
2: for i = 1,. . .,n do
3: if b_i > 0 then
4: R ← RU{i}
5: Fi ← φ
6: while B > 0 do
7: i ← SELECTAGENT(R,$R_1, \ldots,R_n$)
8: f ← CREATEFAKEOBJECT(R_i, F_i, $cond_i$)
9: R_i ← R_i U {f}
10: F_i ← F_i U {f}
11: b_i ← b_i - 1
12: if b_i = 0 then
13: R ← R\{R_i}
14: B ← B − 1

Algorithm 2. Agent Selection for e-random

1: function SELECTAGENT (R,R1, . . .,Rn)
2: i ← select at random an agent from R
3: return i

Explicit Data Request with e-optimal

Still to improve the algorithm for allocation explicit data request we are combining this algorithm with the agent selection for e-optimal method. This algorithm based on e-optimal makes a greedy choice by selecting the agent that will yield the greatest improvement in the sum-objective. The cost of this greedy choice is $O(n^{2})$ in every iteration. The overall running time of e-optimal is $O(n + n^2B) = O(n^2B)$. Here the optimal value is calculated between the different agents and that value is used to attach the fake records in the agent. For calculating the optimal value the co-occurrences between the agents is calculated and the agent with high value is attached with the fake record.

Algorithm 3. Agent Selection for e-optimal

1: function SELECTAGENT (R,R_1, . . .,R_n)

$$i \leftarrow \underset{i':R_{i'}\in \mathbf{R}}{\operatorname{argmax}} \left(\frac{1}{|R_{i'}|} - \frac{1}{|R_{i'}|+1} \right) \sum_{j} |R_{i'} \cap R_j|$$

2: return i

B. Sample data request
Sample Data Request with s-random

Here in this module we present the sample data request with s-random. Here in this method we present the object selection for s-random. In s-random, we introduce vector $a \in N^{|T|}$ that shows the object sharing distribution. In particular, element $a[k]$ shows the number of agents who receive object t_k. Algorithm s-random allocates objects to agents in a round-robin fashion. After the initialization of vectors d and a, the main loop is executed while there are still data objects (remaining > 0) to be allocated to agents. In each iteration of this loop, the algorithm uses function SELECTOBJECT()to find a randomobject to allocate to agent Ui. This loop iterates over all agents who have not received the number of data objects they have requested.

The running time of the algorithm is $O(\tau \sum_{i=1}^{n} m_i)$ and depends on the running time τ of the object selection function SELECTOBJECT(). In case of random selection, we can have $\tau = O(1)$ by keeping in memory a set $\{k' \mid t_{k'} \notin R_i\}$ for each agent U_i.

Algorithm 4. Allocation for Sample Data Requests (SF)

Input: $m_1, . . .,m_n, |T|$. Assuming $m_i \leq |T|$
Output: $R_1,. . .,R_n$
1: a ← $O_{|T|}$. a[k]:number of agents who have
   received object t_k

2: $R_1 \leftarrow \phi , . . .,R_n \leftarrow \phi,$
3: remaining ← $\sum_{i=1} m_i$
4: while remaining > 0 do
5: for all i = 1,. . ., n : $|R_i| < m_i$ do
6: k ← SELECTOBJECT(i,R_i) . May also use additional parameters
7: $R_i \leftarrow R_i$ U $\{t_k\}$
8: a[k] ← a[k]+ 1
9: remaining ← remaining -1

Algorithm 5. Object Selection for s-random

1: function SELECTOBJECT(I,R_i)
2: k ← select at random an element from set $\{k^1|t_k^1 \notin R_i\}$
3: return k

Sample Data Request with s-overlap

In the previous section the distributor can minimize both objectives by allocating distinct sets to all three agents. Such an optimal allocation is possible, since agents request in total fewer objects than the distributor has. This is overcome by presenting an object selection approach for s-overlap. Here in each iteration of allocating sample data request algorithm, we provide agent U*i* with an object that has been given to the smallest number of agents. So, if agents ask for fewer objects than jTj, agent selection for s-optimal algorithm will return in every iterations an object that no agent has received so far. Thus, every agent will receive a data set with objects that no other agent has. The running time of this algorithm is O(1).

Algorithm 6. Object Selection for s-overlap

1: function SELECTOBJECT (i,R_i, a)
2: K ← $\{k \mid k = \operatorname{argmin} a[k^1]\}$
3: k ← select at random an element from set $\{k^1|k^1 \in K \wedge t_k \notin R_i\}$
4: return k

Sample Data Request with s-max

In this module we present an improved algorithm than s-overlap and s-random which we used in allocation algorithm. This algorithm we present here is termed as object selection for s-max. If we apply s-max to the example above, after the first five main loop iterations in algorithm of allocating data request, the R_i sets are:

$R_1 = \{t_1, t_2\}; R_2 = \{t_2\}; R_3 = \{t_3\};$ and $R_4 = \{t_4\}:$

In the next iteration, function SELECTOBJECT() must decide which object to allocate to agent U_2. We see that only objects t_3 and t_4 are good candidates, since allocating t_1 to U_2 will yield a full overlap of R_1 and R_2. Function SELECTOBJECT() of s-max returns indeed t_3 or t_4. The running time of SELECTOBJECT() is O(|T|n).

Algorithm 7. Object Selection for s-max

1: function SELECTOBJECT (i,R_1, . . .,R_n;m_1, . . .,m_n)
2: min_overlap ← 1 . the minimum out of the
   maximum relative overlaps that the allocations of

different objects to U_i yield
3: for k € {k^1 | t_k^1 ¢ R_i} do
4: max_rel_ov← 0 . the maximum relative overlap
   between R_i and any set R_j that the allocation of t_k to U_i
   yields
5: for all j =1; . . . ; n : j ≠ i and t_k € R_j do
6: abs_ov ←|Ri ∩R_j |+ 1
7: rel_ov← abs_ov/min (m_i,m_j)
8: max_rel_ov ←Max (max_rel_ov, rel_ov)
9: if max_rel_ov ≤min_overlap then
10: min_overlap←max_rel_ov
11: ret_k← k
12: return ret _k

In this three overlap values are calculated. The values are absolute overlap value, relative overlap value and minimum overlap value. The minimum overlap value is nothing but the columns allocated to them. The absolute overlap value is the unique column between the agents and the relative overlap value is calculated by dividing the absolute value with the minimum value. According to this value the fake objects will be attached. The agents having the minimum overlap value is attached with the fake records.

Finally with the help of these algorithms the probability is calculated and the guilty agent can be identified. The agent having the highest probability to leak the data is considered as the guilty agent.

Symmetric Inference Model

To represent the possible colluding attacks from any agents to the different data allocation strategies, Here we use the semantic inference model. SIM represents dependent and semantic relationships among attributes of all the entities in the information system. The related attributes (nodes) are connected by three types of relation links: dependency link, schema link, and semantic link.. To evaluate the inference introduced by semantic links, we need to compute the CPT for nodes connected by semantic links.

Symmetric Inference Graph

In order to perform inference at the instance level, we instantiate the SIM with specific entity instances and generate a SIG. Each node in the SIG represents an attribute for a specific instance. The attribute nodes in the SIG have the same CPT as in the SIM because they are just instantiated versions of the attributes in entities. As a result, the SIG represents all the instance-level inference channels.

Instance level dependency link

When a SIM is instantiated, the dependency within- entity is transformed into dependency-within-instance in the SIG. Similarly, the dependency-between-related-entities in the SIM is transformed into a dependency between two attributes in the related instances. This type of dependency is preserved only if two instances

are related by the instantiated schema link. That is, if attribute B in instance *e2* depends on attribute A in instance *e1*, and instances *e1* and *e2* are related by R.

Instance level schema link

The schema links between entities in the SIM represent "key, foreign-key" pairs. At instance level, if the value of the primary key of an instance *e1* is equal to the value of the corresponding foreign key in the other instance *e2* which can be represented as R(*e1*, *e2*), then connecting these two attributes will represent the schema link at the instance level. Otherwise, these two attributes are not connected.

Instance level semantic link

At the instance level, assigning the value of the source node to "unknown" disconnects the semantic link between the attributes of two instances. On the other hand, if two instances have a specific semantic relation, then the inference probability of the target node will be computed based on its CPT and the value of the source node.

IV. EXPERIMENTAL RESULTS EXPLICIT REQUESTS

The goal of these experiments was to see whether fake objects in the distributed data sets yield significant improvement in our chances of detecting a guilty agent. Next, we wanted to evaluate our e-optimal algorithm realative to a random allocation.

It focuses on the scenarios with a few objects that are shared among multiple agents. These are the most interesting scenarios, since object sharing makes it difficult to distinguish a guilty from non-guilty agents. Scenarios with more objects to distribute or scenarios with objects shared among fewer agents are obviously easier to handle.

In our scenarios we have a set of |T|= 10 objects for which there are requests by n= 10 different agents. We assume that each agent request 8 particular objects out these 10. Such scenarios yield very similar agent guilt probabilities and it is important to add fake objects. We generated a random scenario that yielded Δ = 0.073 and min Δ = 0.35 and we applied the algorithms e-random and e-optimal to distribute fake objects to the agents. We varied the number B of distributed fake objects from 2 to 20 and for each value of B we ran both algorithms to allocate the fake objects to agents. We ran e-optimal once for each value of B, since it is a deterministic algorithm. Algorithm e-random is randomized and we ran it 10 times for each value of B. The results we present are the average over the 10 runs.

In Figure 2(a) it shows how fake object allocation can affect negotiation of Δ . Three curves are plotted in the figure. The solid curve is constant and it shows the negotiation of Δ value for an allocation without fake objects.The other two curves look at algorithms e-optimal and e-random. The y-axis shows negotiation of Δ and the x-axis shows the ratio of the number of distributed fake

objects to the total number of objects that the agents explicitly request.

We observe that distributing fake objects can significantly improve on average the chances of detecting a

guilty agent. Even the random allocation of approximately 10% to 15% fake objects yields negotiation of $\Delta > 0.3$. The use of e-optimal improves Δ further, since the e-optimal curve is consistently over the 95% confidence intervals of e-random. The performance difference between the two algorithms would be greater if the agents did not request the same number of objects, since this symmetry allows non-smart fake object allocations to be more effective than in asymmetirc scenarios.
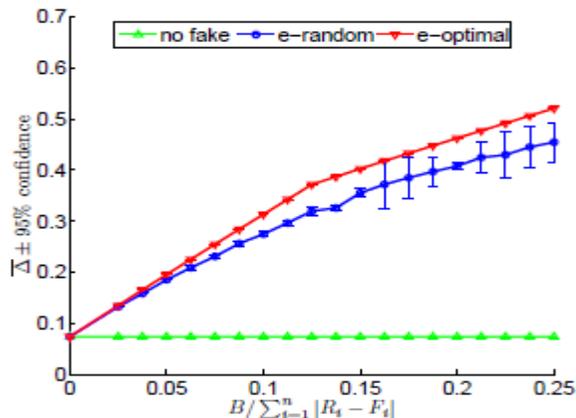
(a) Average $\overline{\Delta}$
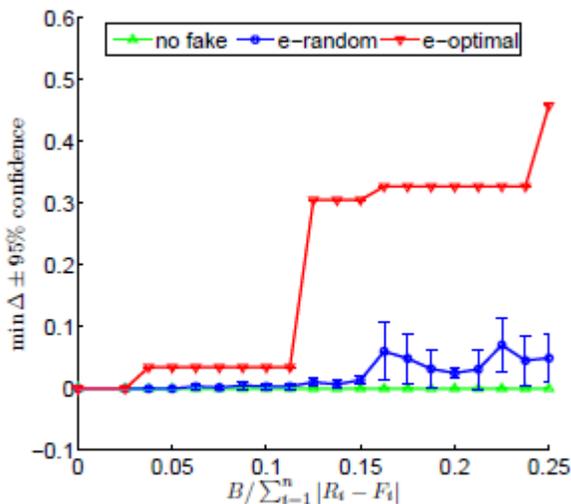
(b) Average $\min \Delta$

Fig 2 . Evaluation of Explicit Data Request Algorithm

Figure 2(b) shows the value of min Δ, as a function of e-optimal the fraction of fake objects. The plot shows that random allocation will yield an insignificant improvement in our chances of detecting a guilty agent in the worst case scenarios. This was expected, since e-

random does not take into consideration which agents must receive a fake object to differentiate their requests from other agents.

Incidentally, the two jumps in the e-optimal curve are due to the symmetry of our scenario. Algorithm e-optimal allocates almost one fake object before allocating a second fake object to one of them.

The result confirms that fake objects can have a significant impact on our chances of detecting a guilty agent. Hence, the performance of e-optimal indicates that our apporximation is effective.

V. CONCLUSION

While transforming the data to the distributors the senders apply perturbation and watermarking techniques to secure them. The approaches used now in the data transformation are not able to detect the data leakage in an efficient way and also they were all restricted or made impossible to satisfy agent's request. They also not help much in finding the guilty agent who leaked the data. The originality of the data and the quality is mainly concentrated here. The techniques cannot be used to some of the sensitive data too as it is impossible to transfer or modify some of the data. So, to find the data leakage and to protect the data from the guilty agents who trying to find the original data, a Semantic Inference Model is going to be introduced. This model is proposed in such a way that it prevents the problem of leakage of data and helps in finding the guilty agent.

REFERENCES

[1]   R. Agrawal and J. Kiernan. Watermarking relational databases. In VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases, pages 155–166. VLDB Endowment, 2002.
[2]   P. Bonatti, S. D. C. di Vimercati, and P. Samarati. An algebra for composing access control policies. ACM Trans. Inf. Syst. Secur., 5(1):1–35, 2002.
[3]   P. Buneman, S. Khanna, and W. C. Tan. Why and where: A characterization of data provenance. In J. V. den Bussche and V. Vianu, editors, Database Theory - ICDT 2001, 8th International
[4]   Conference, London, UK, January 4-6, 2001, Proceedings, volume 1973 of Lecture Notes in Computer Science, pages 316–330. Springer, 2001.
[5]   P. Buneman and W.-C. Tan. Provenance in databases. In SIGMOD '07: Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pages 1171–1173, New York, NY, USA, 2007. ACM.
[6]   Y. Cui and J. Widom. Lineage tracing for general data warehouse transformations. In The VLDB Journal, pages 471–480, 2001.
[7]   S. Czerwinski, R. Fromm, and T. Hodes. Digital music distribution and audio watermarking.
[8]   F. Guo, J. Wang, Z. Zhang, X. Ye, and D. Li. Information Security Applications, pages 138–149. Springer, Berlin / Heidelberg, 2006. An Improved Algorithm to Watermark Numeric Relational Data.
[9]   F. Hartung and B. Girod. Watermarking of uncompressed and compressed video. Signal Processing, 66(3):283–301, 1998.
[10]   S. Jajodia, P. Samarati, M. L. Sapino, and V. S. Subrahmanian. Flexible support for multiple access control policies. ACM Trans. Database Syst., 26(2):214–260, 2001.
[11]   Y. Li, V. Swarup, and S. Jajodia. Fingerprinting relational databases: Schemes and specialties. IEEE Transactions on Dependableand Secure Computing, 02(1):34–45, 2005.
[12]   B. Mungamuru and H. Garcia-Molina. Privacy, preservation and performance: The 3 p's of distributed data management. Technicalreport, Stanford University, 2008.

A Unified Approach on Recommendations in Social Tagging Systems By Limiting Unwanted Request

Manopriya Marimuthu[1], Ms.R.P.Narmadha[2]

[1](M.E Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore) India
[2](Assistant Professor, Department of Computer Science and Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore) India

ABSTRACT

Social Tagging System is the process where user (u) makes their interest by tagging (t) on a particular item (i).These STS are in associated with web 2.0 and has been sourceful information for the users. It provides different types of recommendation in contrast to the current recommendation algorithm, Collaborative Filtering (CF) which apply to two – dimensional data. These data are modeled by a 3-order tensor, on which multiway latent semantic analysis and dimensionality reduction is performed using both the Higher Order Singular Value Decomposition (HOSVD) method and the Kernel-SVD smoothing technique. We provide the 4-order tensor approach, which we named as Tensor Reduction. In particular, the tensor equivalently represents a quadruplet. And also can improve the social tagging efficiency by which unwanted request has been controlled. The results show significant improvements in terms of effectiveness.

KEYWORDS:- SOCIAL TAGS, RECOMMENDER SYSTEMS, TENSORS, HOSVD.

1. INTRODUCTION

Social tagging is the process by which many users add metadata in the form of keywords, to annotate and categorize songs, pictures, products, etc. Social tagging is associated to the "web2.0" technologies and has already become an important source of information for recommender systems. For example, music recommender systems such as Last.fm and MyStrands allow users to tag artist, songs, or albums. In e-commerce sites such as Amazon, users tag products to easily discover common interests with other users. Moreover, social media sites, such as Flickr and YouTube use tags for annotating their content. All these systems can further exploit these social tags to improve the search mechanisms and to personalize the user recommendations. Social tags carry useful information not only about the items they label, but also about the users who tagged. Thus, social tags are a powerful mechanism that reveals three-dimensional correlations between users, tags, and items. Several social tagging systems (STSs), e.g., Last.fm, Amazon, YouTube, etc., recommend items to users, based on tags they have in common with other similar users. Traditional recommender systems use techniques such as Collaborative Filtering (CF) [9], which apply to two-dimensional data, i.e., users and items. Thus, such systems do not capture the multimodal use of tags. To alleviate this problem, Tso-Sutter et al. [13] propose a generic method that allows tags to be incorporated to standard CF algorithms, by reducing the three-dimensional correlations to three 2D correlations and then applying a fusion method to reassociate these correlations.

Another type of recommendation in STSs, e.g., Facebook, Amazon, etc., is to recommend tags to users, based on what tags other users have provided for the same items. Tag recommendations can expose different facets of an information item and relieve users from the obnoxious task to come up with a good set of tags. Thus, tag recommendation can reduce the problem of data sparsity in STSs, which results by the unwillingness of users to provide an adequate number of tags. Recently, several algorithms have been proposed for tag recommendation [4], [5], which project the three-dimensional correlations to three 2D correlations. Then, the two-dimensional correlations are used to build conceptual structures similar to hyperlink structures that are used by Web search engines.

A third type of recommendation that can be provided by STSs is to recommend interesting users to a target user, opting in connecting people with common interests and encouraging people to contribute and share more content. With the term interesting users, we mean those users who have similar profile with the target user. If a set of tags is frequently used by many users, then these users spontaneously form a group of users with common interests, even though they may not have any physical or online connections. The tags represent the commonly interested Web contents to this user group. For example, Amazon recommends to a user who used a specific tag, other new users considering them as interesting ones. Amazon ranks them based on how frequent they used the specific tag.

2. AN OVERVIEW OF RELATED WORK

In this section, we briefly present some of the research literature related to Social Tagging. We also present related work in tag, item, and users recommendation algorithms. Finally, we present works that applied HOSVD in various research domains.Social Tagging is the process by which many users add metadata in the form of keywords to share content. So far, the literature has studied the strengths and the weaknesses of STSs. In particular, Golder and Huberman [12] analyzed the structure of collaborative tagging systems as well as their dynamical aspects. Moreover, Halpin et al. [3] produced a generative model of collaborative tagging in order to understand the dynamics behind it. They claimed that there are three main entities in any tagging system: users, items, and tags.

In the area of item recommendations, many recommender systems already use CF to recommend items based on preferences of similar users, by exploiting a two-way relation of users and items [9]. In 2001, Item-based algorithm was proposed, which is based on the items' similarities for a neighborhood generation. However, because of the ternary relational nature of Social Tagging, two-way CF cannot be applied directly, unless the ternary relation is reduced to a lower dimensional space. Jaschke et al. [11], in order to apply CF in Social Tagging, considered for the ternary relation of users, items, and tags two alternative two-dimensional projections. These projections preserve the user information, and lead to log-based like recommender systems based on occurrence or nonoccurrence of items, or tags, respectively, with the users. Another recently proposed state-of-the-art item recommendation algorithm is tag-aware Fusion [13]. They propose a generic method that allows tags to be incorporated to standard CF algorithms, by reducing the three-dimensional correlations to three 2D correlations and then applying a fusion method to reassociate these correlations.

In the area of tag recommendation, there are algorithms which are based on conceptual structures similar to the hyperlink structures used in Search Engines. For example, Collaborative Tag Suggestions algorithm [5], also known as Penalty-Reward algorithm (PR), uses an authority score for each user. The authority score measures how well each user has tagged in the past. This authority score can be computed via an iterative algorithm similar to HITS. Moreover, the PR algorithm "rewards" the high correlation among tags, whereas it "penalizes" the overlap of concepts among the recommended tags to allow high coverage of multiple factes for anitem. Another state-of-the-art tag recommendation algorithms FolkRank[4]. FolkRank exploits the conceptual structures created by people inside the STSs. Their method is inspired by the seminal PageRank [10] algorithm, which reflects the idea that a web page is important, if there are many pages linking to it, and if those pages are important themselves.

FolkRank employs the same underlying principle for Web Search and Ranking in Social Tagging. The key idea of FolkRank algorithm is that an item which is tagged with important tags by important users becomes important itself. The same holds for tags and users: thus, we have a tripartite graph of vertices which mutually reinforcing each other by spreading their weights. FolkRank is like the Personalized PageRank, which is a modification of global PageRank, and was first proposed for Personalized Web searchind[10].Finally, Xu et.al [6] proposed a method that recommends tags by using HOSVD. Hoever, their method does not cover all three type of recommendation in STSs and misses the comparison with the state-of-art algorithms. In contrast,the approach proposes a unified framework for all recommendation types in STSs. We also combine HOSVD with Kernal-SVD to handle data sparsity,attaining significant improvements in accuracy of recommendation in comparison with simple HOSVD,as will be shown experimentally. In area of discovering shared interest in social networks there are two kind of existing approaches [1]. One is user-centric,which focuses on detecting social interest based on the online connection among users; the other is object-centric, which detects common interest based on common objects fetched by users in a social community.

In the user-centric approach, it can be analyzed for user's online connection and interaction to discover users with particular interest of the given user. Different from this of approach, we aim to find the people who share the same interest no matter whether ther are connected by a social graph or not. In the object-centric approach it explores the common interest among the users based on the common items they fetched in peer-to-peer networks. However, they cannot differentiate the various social interests on the same items due to the fact that users may have different interest for an information item and an item may have multiple facets. In contrast, the approach focuses on directly detecting social interest and recommending users by taking advantage of social tagging, by utilizing users tags. Differently from existing approaches, the method develops a unified framework to concurrently model all three dimensions. Usage data are modeled by a 3-order tensor, on which latent semantic analysis is performed using the HOSVD. Moreover, to address the sparseness problem, we propose the combination of kernel-SVD [7], [8] with HOSVD, which substantially improve the accuracy of item and tag recommendation. HOSVD is a generalization of singular value decomposition (SVD) anh has been successfully applied in several areas. In particular, Wang and Ahuja [2] present a novel multilinear algerbra-based approach to reduced dimensionality representation of multidimensional data,such as image ensembles, video sequences, and volume data. Howeever, they transform the initial tensor (through Clique Expansion algorithm) into lower dimensional spaces, so that clustering algorithm (such as k-means) can be applied.

3. THE PRELIMINARIES

In this section, we summarize the HOSVD procedure.

SVD: The SVD of a matrix $F_{I1 \times I2}$ can be written as a product of three matrices, as shown in equation (1):

$$F_{I1 \times I2} = U_{I1 \times I1} \cdot S_{I1 \times I2} \cdot V^T_{I2 \times I2} ; \qquad (1)$$

Where U is the matrix with the left singular vectors of F, V^T is the transpose of the matrix V with the right singular vectors of F, and S is the diagonal matrix of (ordered) singular values of F.

Tensors: A tensor is a multidimensional matrix. An N-order tensor A is denoted as $A \in R^{I1...IN}$, with elements $a_{i1,...,iN}$. In this paper, for the purposes of the approach, we only use 3-order tensors.

HOSVD: The high-order singular value decomposition generalizes the SVD computation to multidimensional matrices. To apply HOSVD on a 3-order tensor A, three matrix unfolding operations are defined as follows:

$$A1 \in R^{I1 \times I2I3}; \quad A2 \in R^{I2 \times I1I3}; \quad A3 \in R^{I1I2 \times I3} ,$$
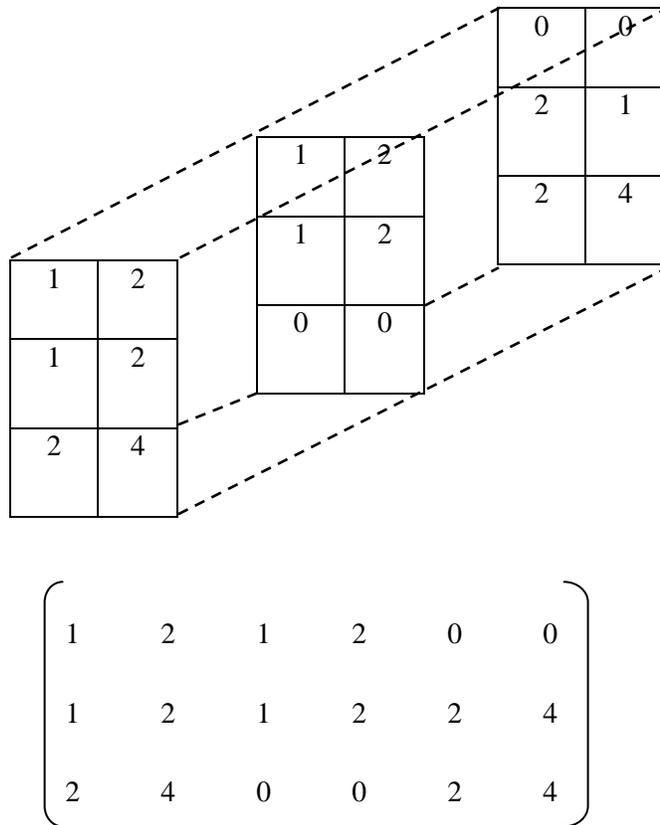
Fig.1. 3 2D matrices

Fig.1: An example tensor A and its 1-mode matrix unfolding A1, where A1, A2, and A3 are called the 1-mode, 2-mode, and 3-mode matrix unfolding of A, respectively. Each A_n ,$1 \leq n \leq 3$, is called the n-mode matrix unfolding of A and is computed by arranging the corresponding fibers of A as columns of An. The left part of Fig. 1 depicts an example tensor, whereas the right part its 1-mode matrix unfolding $A1 \in R^{I1 \times I2I3}$, where the columns (1-mode fibers) of A are being arranged as columns of A1.

4. THE PROPOSED APPROACH

We first provide the outline of the approach, which we name Tensor Reduction, through a motivating example. In this section, we elaborate on how HOSVD is applied on tensors and on how the recommendation of items is performed according to the detected latent associations. Note that a similar approach is followed for the tag and user recommendations.

Three models of Tagging:

After registering a free account, Tagged users can customize their profile page, to which they can post a biography about themselves and their interests, post status updates to inform their friends of their whereabouts and actions, upload photos and albums, and send and receive messages from other users. Model: 1 which represents the user and tag representation in which Fig. 2, represents tagging an item contains a text field, which describes about a particular thing in an item. Item contains photos, videos, etc through which tagging is made.
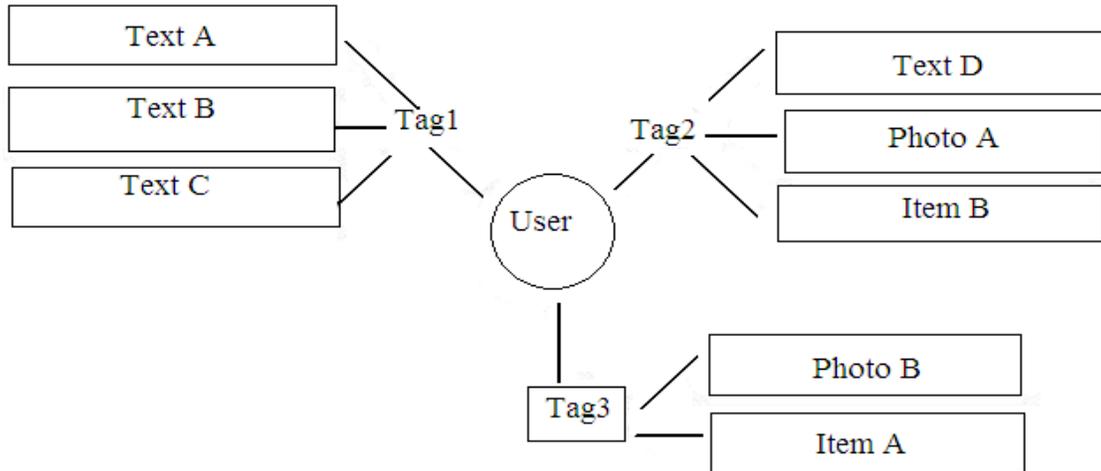
Fig. 2, Model: 1 User – Tag information

The second model fig.3, portrays users that are connected together through their use of tags. This is where real social networking comes in, as users are tagging to relate their concept of information to another user's concept of some piece of information. It may be used inconsistently, by tagging in order that other users see desired information despite the fact that the information is not really classified under their expected concept of that tag.
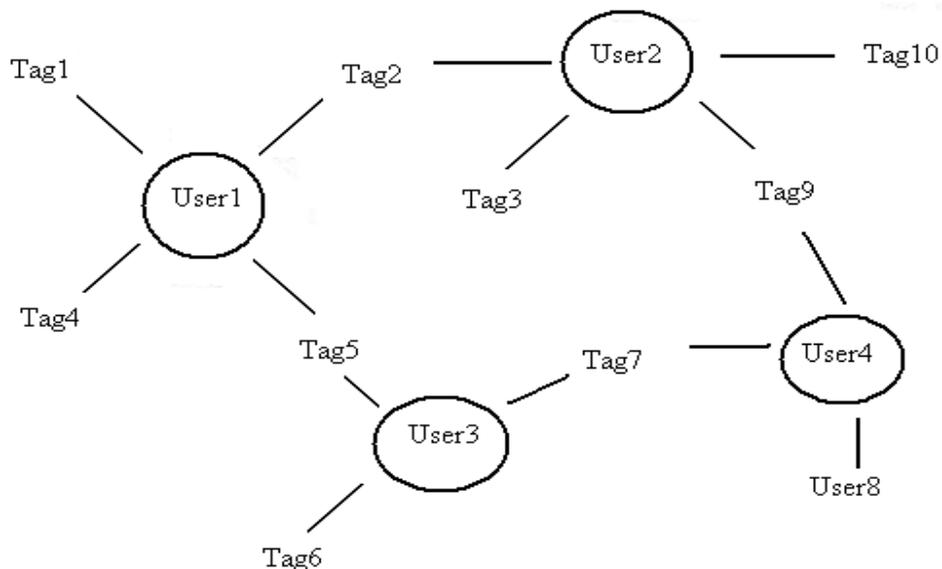
Fig. 3, Model: 2 User - Tag - User

Within the final model, in fig. 4, tags are used to link banks of data (or information) to other information. The tags are acting as metadata to allow search engines to know which information is related to other information. This type of tagging is used greatly within ontologies.
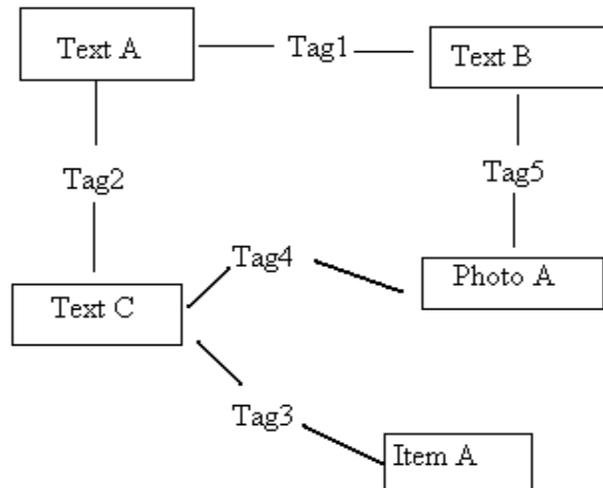
Fig. 4, Model:3 Information - Tag - Information

When using a social tagging systems, to be able to retrieve the information items easily, a user u tags an item i with a tag t. After some time of usage, the tagging system accumulates a collection of usage data, which can be represented by a set of triples {u; i; t}. The Tensor Reduction approach applies HOSVD on the 3-order tensor constructed from these usage data. In accordance with the HOSVD technique introduced, the Tensor Reduction algorithm uses as input the usage data of A and outputs the reconstructed tensor A^. A^ measures the associations among the users, items, and tags. Each element of A^ can be represented by a quadruplet {u; i; t; p}, where p measures the likeliness that user u will tag item i with tag t. Therefore, items can be recommended to u according to their weights associated with {u; t} pair. In this section, in order to illustrate how the approach works, we apply the Tensor Reduction algorithm to a running example. As illustrated in Fig. 5, three users tagged three different items (Web links).

In Fig. 5, the part of an arrow line (sequence of arrows with the same annotation) between a user and an item represents that the user tagged the corresponding item, and the part between an item and a tag indicates that the user tagged this item with the corresponding tag. Thus, the annotated numbers on the arrow lines gives the correspondence between the three types of objects. For example, user U1 tagged item I1 with tag "BMW," denoted as T1. The remaining tags are "Jaguar," denoted as T2, and "CAT," denoted as T3. From Fig. 5, we can see that users U1 and U2 have common interests on cars, while user U3 is interested in cats. A 3-order tensor $A \ 2 \ R^{3 \times 33}$, can be constructed from the usage data.
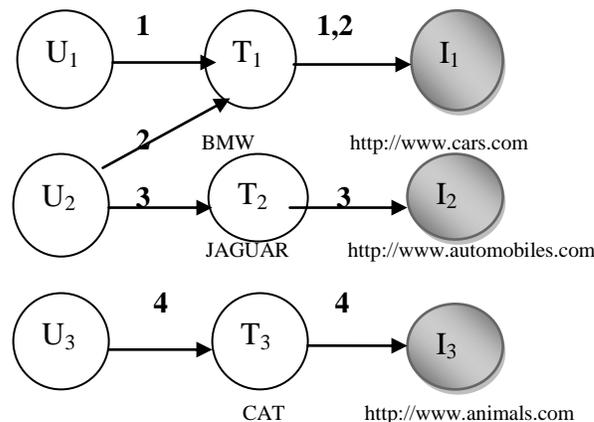
Fig. 5, Usage data of the running example.

We use the co-occurrence frequency (denoted as weight) of each triplet user, item, and tag as the elements of tensor A, which are given in Table 1. Note that all associated weights are initialized to 1. After performing the Tensor Reduction analysis (details of how to do this are given in the following section), we can get reconstructed tensor

TABLE 1
The Elements of the Example Tensor

Arrow Line	User	Item	Tag	Weight
1	U_1	I_1	T_1	1
2	U_2	I_1	T_1	1
3	U_2	I_2	T_2	1
4	U_3	I_3	T_3	1

TABLE 2
The Elements of the Reconstructed Tensor

Arrow Line	User	Item	Tag	Weight
1	U_1	I_1	T_1	0.72
2	U_2	I_1	T_1	1.17
3	U_2	I_2	T_2	0.72
4	U_3	I_3	T_3	1
5	**U_1**	**I_2**	**T_2**	**0.44**

of A^, which is presented in Table 2, whereas Fig. 6 depicts the contents of A^ graphically (the weights are omitted). As shown in Table 2 and Fig. 6, the output of the Tensor Reduction algorithm for the running example is interesting, because a new association among these objects is revealed. The new association is between U1; I2, and T2. This association is represented with the last (bold faced) row in Table 2 and with the dashed arrow line in Fig. 6). If we have to recommend to U1 an item for tag T2, then there is no direct indication for this task in the original tensor A. However, we see that in Table 2 the element recommend the item I2 to user U1, who used tag T2.The resulting recommendation is reasonable, because U1 is interested in
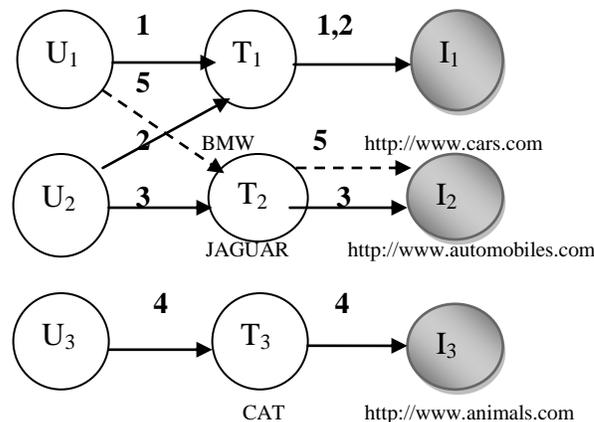
Fig. 6, Illustration of the Tensor Reduction Algorithm output for the running example.

cars rather than cats. That is, the Tensor Reduction approach is able to capture the latent associations among the multitype data objects: user, item, and tags. The associations can then be used to improve the item recommendation procedure, as will be verified by the experimental results. Moreover, for purposes of tag recommendations, we can view the tensor from a different perspective. In particular, the tensor equivalently represents a quadruplet {u, i, t, p} where p is the likeliness that user u will tag item i with tag t. Therefore, tags can be recommended to u according to their weights associated with {u, i} pair. In the running example, if user U1 is about to tag I2, he will be recommended tag T2. Finally, for recommending users, the tensor can be viewed as a quadruplet {t, I, u, p}, where p is the likeliness that tag t will be used to label item i by the user u. Therefore, new users can be recommended for a tag t, according to their total weight, which results by aggregating all items, which are labeled with the same tag by the target user. In the running example, if user U1 tagged item I2 with tag T2, he would receive user U2 as user recommendation.

5. CONCLUSION

Social tagging systems provide recommendations to users based on what tags other users have used on items. In this paper, we developed a unified framework to model the three types of entities that exist in a social tagging system: users, items, and tags. We examined multiway analysis on data modeled as 3-order tensor, to reveal the latent semantic associations between user, items and tags. The multiway latent semantic analysis and   dimensionality reduction is performed by combining the HOSVD method with the Kernel-SVD smoothing technique. The approach improves recommendations by capturing user's multimodal perception of item/tag/user. Moreover, we study a problem of how to provide user recommendation which can have significant applications in real systems but which have not been studied in depth so far in related research. We also performed experimental comparison of the proposed method against state-of-the-art recommendations algorithms, with two real data sets.

The results can show significant improvements in terms of effectiveness measured through recall/precision. As future work, we intend to examine different methods for extending SVD to high-order tensors such as the Parallel Factor Analysis. We also indent to apply different weighting methods for the initial construction of a tensor. A different weighting policy for the tensor's initial values could improve the overall performance of the approach and also limiting the requests.

REFERENCES

[1]   X. Li, L. Guo, and Y. Zhao, "Tag-Based Social Interest Discovery," Proc. ACM World Wide Web (WWW) Conf., 2008.
[2]   H. Wang and N. Ahuja, "A Tensor Approximation Approach to Dimensionality Reduction," Int'l J. Computer Vision, vol. 76, no. 3,pp. 217-229, 2008.
[3]   H. Halpin, V. Robu, and H. Shepherd, "The Complex Dynamics of Collaborative Tagging," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 211-220, 2007.
[4]   A. Hotho, R. Jaschke, C. Schmitz, and G. Stumme, "Information Retrieval in Folksonomies: Search and Ranking," The Semantic Web: Research and Applications, pp. 411-426. Springer, 2006.
[5]   Z. Xu, Y. Fu, J. Mao, and D. Su, "Towards the Semantic Web: Collaborative Tag Suggestions," Proc. Collaborative Web Tagging Workshop at World Wide Web (WWW '06), 2006.
[6]   Y. Xu, L. Zhang, and W. Liu, "Cubic Analysis of Social Bookmarking for Personalized Recommendation," Frontiers of WWW Research and Development—APWeb '06, pp. 733-738. Springer, 2006.
[7]   T. Chin, K. Schindler, and D. Suter, "Incremental Kernel SVD for Face Recognition with Image Sets," Proc. Int'l Conf. Automatic Face and Gesture Recognition (FGR), pp. 461-466, 2006.
[8]   N. Cristianini and J. Shawe-Taylor, Kernel Methods for Pattern Analysis. Cambridge Univ. Press, 2004.
[9]   J. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. Conf. Uncertainty in Artificial Intelligence, pp. 43-52, 1998.
[10] L. Page, S. Brin, R. Motwani, and T. Winograd, "The Pagerank Citation Ranking: Bringing Order to the Web," technical report, 1998.
[11] R. Jaschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G.Stumme, "Tag Recommendations in Folksonomies," Proc. Knowledge Discovery in Databases (PKDD '07), pp. 506-514.
[12] S. Golder and B. Huberman, "The Structure of Collaborative Tagging Systems," technical report, 2005.
[13] K Tso-Sutter, B. Marinho, and L.Schmidt- Thieme, "Tag-Aware Recommender Systems by Fusion of Collaborative Filtering Algorithms," Poc. ACM Symp. Applied Computing(SAC) conf, 2008.

Authors

Ms.Manopriya Marimuthu received B.E degree in CSE from Anna University, Chennai and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. Her research interest includes Computer Networks and Data Mining.

Ms.R.P.Narmadha received the B.E degree in Computer Science and Engineering in Kumaraguru College of Engeering and Technology under Bharathiar University, received the M.E degree in IT in Sathyabama University, Chennai and received the M.B.A degree under Bharathiar University. She is currently working as Assistant Professor in Department of CSE in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Her teaching experience is of 11½ year. Her main research interest is Software Engineering, Data Mining and Image Processing.

Analysis of Various Parameters of Filters (Wavelets) with Curvelet Transform for Denoising in Ultrasound Images

Er. Monica Goyal[1], Er. Sumeet Kaur[2]

*(Department of Computer Science Engineering, GGSCET/ Guru Kashi University,Talwandi Sabo) India
** (Department of Computer Science Engineering, YOCE/Punjabi University,Talwandi Sabo) India

ABSTRACT

Ultrasonography is considered to be one of the most powerful techniques for imaging organs and soft tissue structures in human body. It is preferred over other medical imaging methods because it is non-invasive, portable, and versatile and does not use ionising radiations. Despite their obvious advantages, ultrasound (US) images are contaminated with multiplicative noise called 'speckle' which is one of the major sources of image quality degradation. In the medical literature, speckle has been treated as a distracting artifact as it tends to degrade the resolution and the object detectability. Image denoising is used to remove the noise while retaining as much as possible the important signal features. The purpose of image denoising is to estimate the original image form the noisy data. Image denoising is still remains the challenge for researchers because noise removal introduces artifacts and causes blurring of the images.

Keywords: Speckle noise, Denoising, Simulation, Blurred, Speckle reduction.

1.INTRODUCTION

1.1 Ultrasound Images
Ultrasonography is considered to be one of the most powerful techniques for imaging organs and soft tissue structures in human body.

Despite their obvious advantages, ultrasound(US) images are contaminated with multiplicative noise called 'speckle' which is one of the major sources of image quality degradation.

In the medical literature, speckle has been treated as a distracting artifact as it tends to   degrade the resolution and the object detectability. Moreover, in US images the speckle noise has a spatial correlation length on each axis, which is same as resolution cell size. This spatial correlation makes the speckle suppression a very difficult and delicate task, hence, a trade-off has to be made between the degree of speckle suppression and feature preservation.

1.2 Speckle Noise
Speckle significantly degrades the image quality and hence, makes it more difficult for the observer to discriminate fine detail of the images in diagnostic examinations. Speckle is a form of multiplicative noise, which makes visual interpretation difficult. Laser holography and ultrasound imaging are two techniques susceptive to speckle degradation. Speckle noise causing greater degradation within bright areas of an image than in dark areas.

1.3 Image Denoising
Image denoising is used to remove the noise while retaining as much as possible the important signal features. The purpose of image denoising is to estimate the original image form the noisy data. Image denoising is still remains the challenge for researchers because noise removal introduces artifacts and causes blurring of the images.

2. LITERATURE SURVEY
A new multiscale non-linear method for speckle suppression in ultrasound images is presented. The main innovation is the use of realistic distributions of the wavelet coefficients. By combining these distributions with a simple shrinkage function (soft-thresholding), a closed-form expression for soft thresholding is derived analytically. [1]

A new and efficient technique for despeckling medical US images has been proposed, which relies on the Rayleigh distribution of speckle noise and Gaussian prior for modelling the wavelet coefficients in a logarithmically transformed US image. [2]

????.....This work describes the implementation, testing and evaluation of popular denoising algorithms for the denoising of low-field MR images. [3]

For Medical field denoising, an image prior model was purposed using Markov Random Field. The parameters on the model are learned from PCA and MLE. Based on this model, image denoising can be done by Bayesian analysis. [4]

A simple and subband adaptive threshold is proposed to address the issue of image recovery from its noisy counterpart. It is based on the generalized Guassian distribution modeling of subband coefficients. The image denoising algorithm uses soft thresholding to provide smoothness and better edge preservation at the same time. [5]

A strategy for digitally implement both the ridgelet and the curvelet transforms. Curvelet thresholding rivals sophisticated techniques that have been the object of extensive development over the last decade. [6]

Curvelets provide a powerful tool for representing very general linear symmetric systems of hyperbolic differential equations. [7]

A novel implementation of the discrete curvelet transform is proposed in this work. The transform is based on the Fast Fourier Transform (FFT) and has the same order of complexity as the FFT. The discrete curvelet functions are defined by a parameterized family of smooth windowed functions that are 2-pi periodic and form a partition of unity. The transform is named the Uniform Discrete Curvelet Transform (UDCT) because the centers of the curvelet functions at each resolution are located on a uniform grid. [8]

?????...A new approach for SAR image enhancement and change detection based on the curvelet transform has been proposed and applied to TerraSAR-X data of the city center of Munich. [9]

An adaptive threshold estimation method for image denoising in the curvelet domain by using mean, (Spatial Frequency Measure) SFM and (Difference Operator)DOP, experiment work on Lena, cameraman and boat gray test images at different type of noises (Random, Salt & pepper ,Gaussian ,Speckle) showed that the proposed adaptive threshold method success to estimate and reduce noise from image and it is more effective at reduce noise from image than (Rudin-Osher-Fatemi) ROF filter and Non Local Mean algorithm. The proposed adaptive estimation method introduced betterresults than (Rudin-Osher-Fatemi) ROF filter, Non Local Mean algorithm and Wiener filter at reduce noise (Random, Salt & pepper, Gaussian) according to increasing of PSNR values of enhanced images by 0.044 at Random, 1.05 at salt &pepper and 0.457 at Gaussian noise. [10]

3. TECHNIQUES FOR DENOISING

3.1 Curvelet Transform Techniques
Curvelet Transform is a new multi-scale representation most suitable for objects with curves. Developed by Candès and Donoho (1999).

A discontinuity point affects all the Fourier coefficients in the domain. Hence the FT doesn't handle point's discontinuities well. Using wavelets, it affects only a limited number of coefficients. Hence the WT handles point discontinuities well. Discontinuities across a simple curve affect all the wavelets coefficients on the curve. Hence the WT doesn't handle curves discontinuities well. Curvelets are designed to handle curves using only a small

number of coefficients. Hence the CvT handles curve discontinuities well.

The Curvelet Transform includes four stages:
- Sub-band decomposition
- Smooth partitioning
- Renormalization
- Ridgelet analysis

3.1.1 Algorithm
1. Sub-band decomposition
$$f \mapsto \left(P_0 f, \Delta_1 f, \Delta_2 f, \ldots\right)$$

- Dividing the image into resolution layers.
- Each layer contains details of different frequencies:
- o P_0 – Low-pass filter.
- o $\Delta_1, \Delta_2, \ldots$ – Band-pass (high-pass) filters.
- The original image can be reconstructed from the sub-bands:

$$\|f\|_2^2 = \|P_0 f\|_2^2 + \sum_s \|\Delta_s f\|_2^2$$

Energy preservation

$$f = P_0\left(P_0 f\right) + \sum_s \Delta_s\left(\Delta_s f\right)$$

- Low-pass filter Φ_0 deals with low frequencies near $|\xi| \leq 1$.
- Band-pass filters Ψ_{2s} deals with frequencies near domain $|\xi| \in [2^{2s}, 2^{2s+2}]$.
- Recursive construction – $\Psi_{2s}(x) = 2^{4s}\Psi(2^{2s}x)$.
- The sub-band decomposition is simply applying a convolution operator:

$$P_0 f = \Phi_0 * f \qquad \Delta_s f = \Psi_{2s} * f$$

- The sub-band decomposition can be approximated using the well known wavelet transform:
- o Using wavelet transform, f is decomposed into S_0, D_1, D_2, D_3, etc.
- o $P_0 f$ is partially constructed from S_0 and D_1, and may include also D_2 and D_3.
- o $\Delta_s f$ is constructed from D_{2s} and D_{2s+1}.

2. **Smooth partitioning**
- A grid of dyadic squares is defined:

$$Q_{(s,k_1,k_2)} = \left[\frac{k_1}{2^s}, \frac{k_1+1}{2^s}\right] \times \left[\frac{k_2}{2^s}, \frac{k_2+1}{2^s}\right] \in \mathbf{Q}_s$$

- $\mathbf{Q}_s$ – All the dyadic squares of the grid.
- Let w be a smooth windowing function with 'main' support of size $2^{-s} \times 2^{-s}$.
- For each square, w_Q is a displacement of w localized near Q.
- Multiplying $\Delta_s f$ with w_Q ($\forall Q \in \mathbf{Q}_s$) produces a smooth dissection of the function into 'squares'.

- The windowing function w is a nonnegative smooth function.

$$h_Q = w_Q \cdot \Delta_s f$$

- Partition of the energy:
o The energy of certain pixel (x_1, x_2) is divided between all sampling windows of the grid.

$$\sum_{k_1,k_2} w^2\left(x_1 - k_1, x_2 - k_2\right) \equiv 1$$

- Reconstruction:

$$\sum_{Q \in \mathbf{Q}_s} w_Q \cdot h_Q = \sum_{Q \in \mathbf{Q}_s} w_Q^2 \cdot h = h$$

- Parserval relation:

$$\sum_{Q \in \mathbf{Q}_s} \left\| h_Q \right\|_2^2 = \sum_{Q \in \mathbf{Q}_s} \int w_Q^2 \cdot h^2 = \int \sum_{Q \in \mathbf{Q}_s} w_Q^2 \cdot h^2 = \int h^2 = \left\| h \right\|_2^2$$

3. Renormalization

- Renormalization is centering each dyadic square to the unit square $[0, 1] \times [0, 1]$.
- For each Q, the operator T_Q is defined as:

$$\left(T_Q f\right)\left(x_1, x_2\right) = 2^s f\left(2^s x_1 - k_1, 2^s x_2 - k_2\right)$$

- Each square is renormalized

$$g_Q = T_Q^{-1} h_Q$$

4. Ridgelet analysis
A) Before ridgelet transform:

- The $\Delta_s f$ layer contains objects with frequencies near domain $|\xi| \in [2^{2s}, 2^{2s+2}]$.
o We expect to find ridges with width $\approx 2^{-2s}$.
- Windowing creates ridges of width $\approx 2^{-2s}$ and length $\approx 2^{-s}$.
- The renormalized ridge has an aspect ratio of width $\approx$ length^2.
- We would like to encode those ridges efficiently
o Using the **Ridgelet Transform**
- Ridgelet are an orthonormal set $\{\rho_\lambda\}$ for $L^2(\Re^2)$
- Divides the frequency domain to dyadic coronae $|\xi| \in [2^s, 2^{s+1}]$.
- In the angular direction, samples the s-the corona at least 2^s times.
- In the radial direction, samples using local wavelets.
- The ridgelet element has a formula in the frequency domain:

where,

$$\hat{\rho}_\lambda\left(\xi\right) = \tfrac{1}{2} |\xi|^{-\frac{1}{2}} \left(\hat{\psi}_{j,k}\left(|\xi|\right) \cdot \omega_{i,l}(\theta) + \hat{\psi}_{j,k}\left(-|\xi|\right) \cdot \omega_{i,l}(\theta + \pi)\right)$$

- $\omega_{i,l}$ are periodic wavelets for $[-\pi, \pi)$.
- i is the angular scale and $1 \in [0, 2^{i-1} - 1]$ is the angular location.
- $\psi_{j,k}$ are Meyer wavelets for $\Re$.
- j is the ridgelet scale and k is the ridgelet location.

Ridgelet transform:
- Each normalized square is analyzed in the ridgelet system:

$$\alpha_{(Q,\lambda)} = \left\langle g_Q, \rho_\lambda \right\rangle$$

o The ridge fragment has an aspect ratio of $2^{-2s} \times 2^{-s}$.
o After the renormalization, it has localized frequency in band $|\xi| \in [2^s, 2^{s+1}]$.
o A ridge fragment needs only a very few ridgelet coefficients to represent it.

3.2 Discrete Wavelet Transform
Discrete Wavelet Transform (DWT) is introduced to overcome the redundancy problem of CWT. The approach is to scale and translate the wavelets in discrete steps.

$$\mathrm{DWT}(\tau_0, s_0) = \frac{1}{\sqrt{s_0^f}} \int_{-\infty}^{\infty} f(t)\, \psi\left(\frac{t - k\tau_0 s_0^f}{s_0^f}\right) dt \qquad (3.1)$$

Where s_0^f is the scaling factor τ_0 is the translating factor, k and j are just integers.
Subsequently, we can represent the mother wavelet in term of scaling and translation of a dyadic transform as

$$\psi_{j,k}(t) = 2^{-f/2} \psi\left(2^{-f} t - k\right) \qquad (3.2)$$

Replacing eqn, the coefficients of DWT can be represented as

$$C_{f,k} = 2^{-f/2} \int_{-\infty}^{\infty} f(t)\, \psi\left(2^{-f} t - k\right) dx \qquad (3.3)$$

By applying DWT, the image is actually divided i.e., decomposed into four sub-bands and critically sub sampled as shown in fig 3.1(1):

LL1	LH1
HL1	HH1

LL2	LH2	
HL2	HH2	LH1
HL1		HH1

(a) One-Level   (b) Two-Level
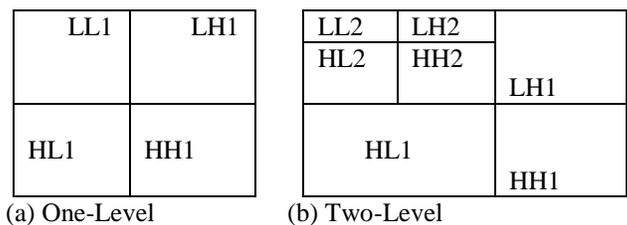
Fig 3.1(1): Image Decomposition

3.3 Denoise Procedure
Wavelets are especially well suited for studying non-stationary signals and the most successful applications of wavelets have been in compression, detection and denoising. In recent years, there has been a fair amount of research on wavelet based image denoising. The algorithm is simple but provides good results for a wide variety of signals. The method consists of applying the DWT to the original data, thresholding the detailed wavelet coefficients and inverse transforming the set of thresholded coefficients

to obtain the denoised signal. This scheme is known as Visu-Shrink and is further illustrating in the block diagram of Figure 3.2.
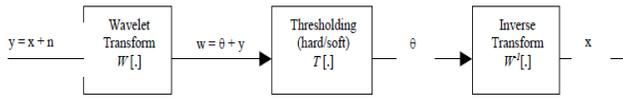
Figure 3.2: Block diagram for DWT based denoising framework

3.4 Image Denoising Filter Methods

The various filter methods for image denoising:

1) Median Filter
2) Wiener filter

Median Filter:
This filter sorts the surrounding pixels value in the window to an orderly set and replaces the center pixel within the define window with the middle value in the set.

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

Wiener Filter:
Wiener2 low pass-filters an intensity image that has been degraded by constant power additive noise. Wiener2 uses a pixel wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel.
J = wiener2 (I, [m n], noise) filters the image I using pixel wise adaptive Wiener filtering, using neighborhoods of size m-by-n to estimate the local image mean and standard deviation. If you omit the [m n] argument, m and n default to 3. The additive noise (Gaussian white noise) power is assumed to be noise.

[J, noise] = wiener2 (I, [m n]) also estimates the additive noise power before doing the filtering. Wiener2 returns this estimate in noise.

3.5 Discrete Algorithm
In this work, the algorithm via the wavelet shrinkage technique is as follows:

1. Take a given original image.
2. Take the logarithmic transform of speckled image.

I. Perform multiscale decomposition of the log transformed image using wavelet transform.
II. Estimate the noise variance d using the below formula[2]

$$\hat{\sigma}^2 = \left[\frac{median(|Y_{ij}|)}{0.6745}\right]^2, \qquad Y_{ij} \in \text{ Subband HH}_1$$

III. For each level in sub bands, compute the scale parameter K using the below formula.[2]

$$K = \sqrt{\log(L_k)}$$

IV. For each subband (except the lowpass residual).Compute the standard deviation **ox** using the below formula.[2]

$$\hat{\sigma}_x \approx \sqrt{\max\left(\hat{\sigma}_y^2 - \hat{\sigma}^2, \, 0\right)}$$

V. Compute threshold TN using below formula[2]

$$T_N = K\frac{\sigma^2}{\sigma_x}$$

if' subband variance **U:** is greater than noise variance, otherwise set TN to maximum coefficient of the sub band.
VI. Apply soft thresholding to the noisy coefficients.
VII. Invert the multiscale decomposition to reconstruct the denoised image **f** ;
VIII. Take the exponential of the 'reconstructed image obtained from step 9[2].

4. PARAMETER METRICS
4.1 PSNR
A high quality image has small value of Peak Signal to Noise Ratio (PSNR).

PSNR is defined as follow:

$$\mathbf{PSNR} = \left[10\log\frac{255^2}{MSE}\right].$$

2. Coefficient of Correlation (CoC)

$$CoC = \frac{\varepsilon(g-\bar{g}).\left(\hat{g}-\bar{\hat{g}}\right)}{\sqrt{\varepsilon\left(\Delta g - \overline{\Delta g}\right)^2.\varepsilon\left(\Delta g - \overline{\Delta \hat{g}}\right)}}$$

Where $\bar{g}$ is mean of original image, $\bar{\hat{g}}$ is mean of denoised image.

5. PERFORMANCE ANALYSIS

a) Original image   b) Noisy image

c) Decomposed structure

d) Wiener filter image

e) Median Filter image

f) Curvelet transform image
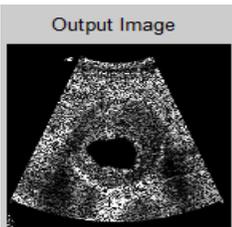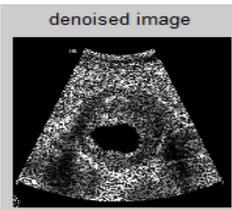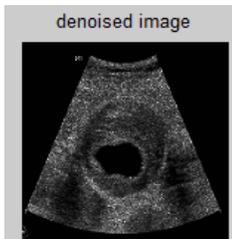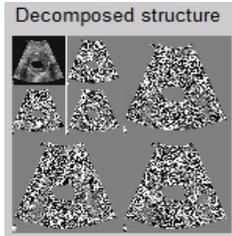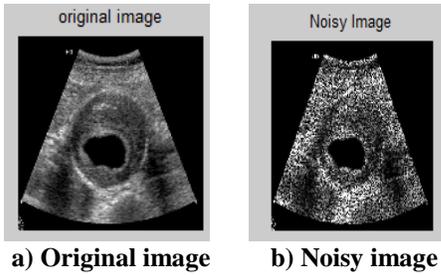
Fig 5.1(1) Images after applying filters ad curvelet transform techniques for denoising.

Table 5.1(1) PSNR results test for discrete wavelet transform (filters) and Curvelet transform on ultrasound image and σ value (1) Curvelet transform (2) wiener filter (3) media filter

Image	Curvelet transform	Wiener filter	Median filter
σ= 0.5	11.6920	21.8049	**23.5206**
σ= 1.0	11.4968	19.7270	**20.7281**
σ =1.5	12.1962	18.6646	**19.2271**
σ= 2.0	12.1746	18.0021	**18.2348**
σ =2.5	15.3100	17.5245	**17.5294**

Table 5.2(2) Coc results test for discrete wavelet transform (filters) and Curvelet transform on ultrasound image and σ value (1) Curvelet transform (2) wiener filter (3) media filter

Image	Curvelet transform	Wiener filter	Median filter
σ= 0.5	0.5485	0.8539	**0.8950**
σ= 1.0	0.5478	0.8082	**0.8297**
σ =1.5	0.5869	0.7848	**0.7852**
σ= 2.0	0.5849	**0.7732**	0.7541
σ =2.5	0.7425	**0.7632**	0.7326

6. CONCLUSION

In comparison of different filtering methods and curvelet transform method, a novel multiscale nonlinear method for speckle suppression in ultrasound images is presented. Experiments are conducted to access the better performance from denoising filtering methods. The result showed in table 5.1(1) and 5.2(2) shows that Filtering Methods of discrete wavelets transform produce better result than curvelet transform methods. Wavelet based denoising algorithms uses soft thresholding to provide smoothness and better edge preservation. Wiener filter removes noise significantly and outperforms the median filter. Despite the significance of Curvelet transform having discontinuities working over arc the wiener filter display images with more clarity.

REFERENCES

[1] Savita Gupta, L.Kaur, R.C. Chauhan and S.C.Saxena, 'A wavelet based Statistical Approach for Speckle Reduction in Medical Ultrasound Images', MBEC Med. Biol, Eng. Comput., March 2004, Vol42, and 189-192.

[2] Lakhwinder Kaur, Savita Gupta and R.C.Chauhan, 'Image denoising using Wavelet Thresholding'.

[3] Ting-Li Chen,'A Markov Random Field Model for Medical Image Denoising', 978-1-4244-4134-1/09/$25.00 ©2009 IEEE.

[4] NF Ishak, MJ Gangeh, R Logeswaran "Comparison of Denoising Techniques Applied on   Low-field MR Brain Images", in 2008.

[5] Jean-Luc Starck, Emmanuel J. Candμes, David L. Donoho "The Curvelet Transform for Image Denoising", in November 15, 2000.

[6] Emmanuel J. Candμes and Laurent Demanet "Curvelets and Wave Equations", in september 11, 2004.

[7] T.T. Nguyen (Ecole des Mines de Paris) & H. Chauris* (Ecole des Mines de Paris) "Uniform Discrete Curvelet Transform for Seismic Processing", in 2008.

[8] Andreas Schmitt, Birgit Wessel, Achim Roth "CURVELET APPROACH FOR SAR IMAGE DENOISING, STRUCTURE ENHANCEMENT, AND CHANGE DETECTION" in September 2009.

[9] Aliaa A.A., Youssif A.A.Darwish, A.M.M.Madbouly "Adaptive Algorithm for Image Denoising Based on Curvelet Threshold", in January 2010

A New OLSR Routing Protocol in Cognitive Wireless Mesh Networks

Venkaiahnaidu.A[1], B.RameshBabu[2], Sk.MansoorRahaman[3], K.RajasekharaRao[4]

[1]Student, Department Of CSE, K.L.University

[2]Assoc.Prof, Department Of CSE, K.L.University

[3]Student, Department Of CSE, K.L.University

[4]Principal, K.L.University

Abstract

Ad hoc On-Demand Distance Vector (AODV) is a reactive routing protocol that establishes a route based on a requirement. By avoiding counting-to-infinity problem it performs better when compared to most common routing protocols of the Internet such as IGRP, EIGRP, OSPF, RIP, and IS-IS. Though reactive and on-demand, AODV cannot capitalize on all available dynamic spectrum resources in cognitive wireless mesh networks yet maintaining a high throughput to route packets. For that reason in our paper we propose an enhanced AODV protocol named AODV-COG that provides an interface to a route for efficient usage of the spectrum and finding a path with high throughput among the paths with same hop count. Simulations in NS-2 require changes to support cognitive wireless mesh network conditions. In this paper we are focusing on Optimized Link State Routing Protocol (OLSR).

Keywords-Cognitive mesh networks, routing protocol, AODV, OLSR.

1. Introduction

Today wireless communication industry is growing very fast. The inconsistency between the limited spectrums of wireless applications becomes more and more important. The resources which are used by the users are not given full licensed. According to the Federal Commission FCC data, the utilization rate of allocated spectrum is only 15% -85%. Cognitive radio technology will manage to resolve the issue. In mesh networks to improve the capacity and network throughput will cognitive radio technology.

In this research, the cognitive wireless mesh networks mainly addresses on the sensing technology and spectrum sharing scheme and the routing for utilization of multiple channels in multi-hop cognitive mesh networks will face several problems. In multiple channel networks, the set of channels available for each node is not static. In this paper, we are using the advanced AODV that uses the routing request and response messages from node to node SOP(Spectrum Opportunities) information will explore the calculated information of the routes. In this paper, we use equal number of interfaces and channels. In this design

will create the channel for each interface. For the cognitive radio condition will use network simulator 2 (NS-2). The protocol which we are using will select the best path. The contribution of the paper

- We modify the Network Simulator 2(NS-2) to support the cognitive function; the readers can also use our architecture to evaluate their own routing protocol for cognitive mesh networks.
- We implement the AODV to increase the throughput and performance in the cognitive mesh networks.

Many researchers in the past have compared the before mentioned protocols considering the standard wireless ad hoc networks. But mesh networks are different from the other networks. Here we are explaining the problems in the mesh networks and solving the problems. One of the critical issues of the mesh networks is scalability. In this paper we are improving the scalability and throughput using OLSR routing protocol to reduce the overhead and then compare the OLSR to AODV in terms of packet delivery ratio.

2. Related Work

After researching many protocols will use the extension for the AODV is OLSR. First we will identify the most popular ad hoc routing protocols that are having direct candidates for the routing protocols in wireless mesh networks. The routing protocol is suitable for small. Large-scale wireless mesh networks. The ad hoc networks traditionally divided into two categories: on-demand (reactive) and table driven (proactive) protocols. In the reactive protocols the route path is finalize only when the node is transferring data packets to send. All the results we will show using NS-2 simulator.

3. The ETX metric

AODV-COG uses the ETX metric in its main two functions, so let us introduce the ETX metric and our own approach to get the parameters those are used to calculate the ETX first. The ETX determine routes with the fewest excepted number of transmissions required to deliver a

packet to its destination, including re-transmissions. It aims to choose routes with high end-to-end throughput. ETX's effectiveness has been proved in [8]. We use the forward and reverse delivery ratios of a link to calculate the ETX. The forward delivery ratio, d_f, is the measured probability that a packet successfully arrives at the destination; the reverse delivery ratio, d_r, is the probability that the ACK packet is successfully received [8].

Our method to get the d_f, d_r and calculate the ETX is as follows:

1). In AODV protocol, the nodes broadcast hello message every HELLO_INTERVAL milliseconds. In our protocol, we set the HELLO_INTERVAL as definite value milliseconds. Then every millisecond, every node sends a hello message through all channels. Because the Hello messages are broadcast, 802.11 protocols do not acknowledge or retransmit them. We define a two-dimensional array *Numhello[i][j]* for every node to record the number of hello messages sent by its neighbor node *i* using channel *j*.

2). Every *t* milliseconds, we calculate the d_r for each node on each channel. We define another two-dimensional array d_r *[i][j]* to record the delivery ratio from the neighbor node *i* through channel *j* :

$$d_r\,[i][j] = \frac{Numhello[i][j]}{m/t}$$

m is the number of hello messages that should have been received during last *t* milliseconds.

3). We can see that the *Numhello[i][j]* stored in node *k* should also be used to calculate d_f *[k][j]* for node i ,so every *t* milliseconds, we send a new packets named hello-ack packets which carry the *Numhello[i][j]* information back to the node *i*. Then the node *i* should calculate the d_f .After the unicast transmit of the hello-ack packets is complete, we clear the array *Numhello[i][j]* for next *t* period.

4). When we get the d_f, d_r, we should calculate the ETX using the following function [8] to calculate the ETX:
$$ETX= \quad 1/ (\; d_f \times d_r \;)$$

4. AODV-COG Protocol
Our routing protocol is based on the Ad-hoc On-Demand Distance Vector (AODV) protocol.. The AODV-COG modifies the AODV protocol mainly in two parts: One is the path selection function, which is used to fine a path with higher throughput. The other is the interface assignment function, which is used to select an interface to route packets.

4.1 Path selection function
The AODV protocol determines the routes using the minimum hop-count metric which is not always finds the path with the largest throughput. Even if the best route is a minimum hop-count route, there may be another route with the same hop counts but achieves larger throughput. Our protocol aims to find the route which gains the highest throughput among the routes with the same minimum hop counts, that is to say, we hope to find a route which is the shortest path route and meanwhile the one with the largest throughput. We combine the ETX and the minimum hop-count in AODV-COG to achieve our goal. We only describe the differences between our protocol and the AODV:

1). We add a new field *rq_sumetx* in the RREQ packet. It is used to record the sum of links' ETX of the path which the RREQ passing through. A new field *route_etx* is also added in the route table. The *rq_sumetx* and *route_etx* are initialized to 0. When a node needs to send a route request, it broadcasts the RREQ on all available channels without dependence on the common control channel.

2). When the intermediate node receive RREQ messages, firstly, it use the function (2) to calculate the ETX and then add the *rq_sumetx* by the result, for instance, when the node *k* receives a RREQ message from node *i* through channel *j* then the
$$rq\_sumetx_{after}= rq\_sumetx_{before}+1/ (dr[i][\,j]\; Xd_f[i][j])$$
Secondly, the node will check the RREQ to determine whether to update the route table. The route is only updated if:
(i)   The sequence number stored in the RREQ is either higher than the destination sequence number in the route table, or
(ii)   The sequence numbers are equal, but the hop count (of the RREQ) is smaller than the hop count in the routing table, or
(iii) The sequence number and the hop count is equal but the *rq_sumetx* is smaller than the *route_etx* in the route table. If the node retransmits the packet, it transmits on all available channels. The information of ETX which is piggybacked by RREQ messages is forwarded in the broadcast process.

3). Like we have done in RREP, we also identify a new field *rp_sumetx* in RREP packet, if the node receives a Route Reply packet (RREP), it updates the *rp_sumetx* in the RREP and then use the same regulation as 2) to update the route table.

5. Optimized Link State Routing Protocol
The information in this section concerning the Optimized Link State Protocol is taken from its RFC 3561. Optimized Link State Protocol (OLSR) is a proactive routing protocol, so the routes are always immediately available when needed. OLSR is an optimization version of a pure

link state protocol. So the topological changes cause the flooding of the topological information to all available hosts in the network. To reduce the possible overhead in the network protocol uses Multipoint Relays (MPR). The idea of MPR is to reduce flooding of broadcasts by reducing the same broadcast in some regions in the network, more details about MPR can be found later in this chapter. Another reduce is to provide the shortest path. The reducing the time interval for the control messages transmission can bring more reactivity. OLSR uses two kinds of the control messages: Hello and Topology Control (TC). Hello messages are used for finding the information about the link status and the host's neighbors. With the Hello message the Multipoint Relay (MPR) Selector set is constructed which describes which neighbors has chosen this host to act as MPR and from this information the host can calculate its own set of the MPRs. the Hello messages are sent only one hop away but the TC messages are broadcasted throughout the entire network. TC messages are used for broadcasting information about own advertised neighbors which includes at least the MPR Selector list.

How the OSLR works in the NS-2
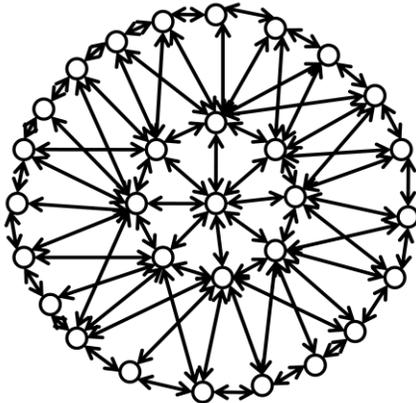
Flooding a packet in a wireless multi-hop network the arrows show all transmissions
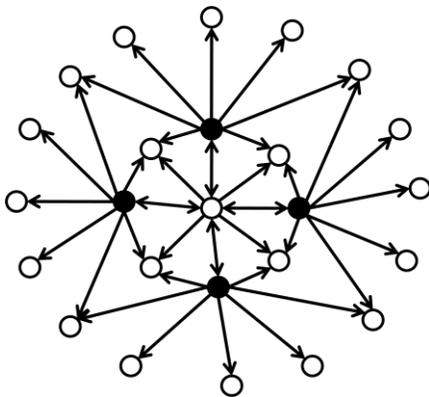
Flooding a packet in a wireless multi-hop network from the center node using MPR's(black) The arrows show all transmissions

6.The Performance Evaluation in NS-2

The verity of needs we are having with Network simulators. Compare with the other containing multiple networked computers, routers and data links, network simulators are fast and inexpensive. Network simulators are particularly use for the researchers to test new networking protocols or changes to existing protocols in a controlled and reproducible environment. The network simulator is a software or hardware that estimates the behavior of a network. NS2 will give the support between networking research and education. We can design the protocol and the study of the traffic in the different types of networks. The protocol that we are using will compare the show the results effectively. We can implement the any of the new architecture designs are also supported which will make easy to implement. NS2 will provide the collaborative environment which will be the freely distributed to all the users, and it is the software which is open source. Using NS2 as the software we can show the result very effectively and accurately the user get the confidence of showing their results.

In this paper we are implementing the AODV and OLSR protocols for the increasing the throughput and efficient of packet transfer. We can easily build the network structure and topology which is just the surface of your simulation. Here we can configure the network parameters very easily.

7. Conclusion

In this Existing we use a novel cognitive routing protocol (AODV-COG). It has two parts: one is the path selection function that combines the min-hop metric and the ETX metric to find the best route and the other is a Channel assignment module that can perceive the dynamic SOPs periodically and change the interface if necessary .We also add a new module named Send Control Module in the NS-2 simulator to support the simulation of cognitive routing protocol. Based on the simulation results, AODV-COG is able to increase the overall throughput in the cognitive mesh network. In this proposed based on the simulation results with OLSR compare with AODV-COG we will able to increase the overall throughput in the cognitive mesh networks.

8. Acknowledgement

9. References

[1] Ian F. Akyildiz, et al, "NeXt generation /dynamic spectrum access/cognitive radio wireless networks: A survey",Computer Networks, May,2006

[2] R. Etkin, A, Parekh, D. Tse, "Spectrum sharing for unlicensed bands " , DySPAN 2005, Nov. 2005

[3] Geng Cheng, Wei Liu, Yunzhao Li, and Wenqing Cheng, "Joint On-demand Routing and Spectrum Assignment in Cognitive Radio Networks" IEEE.ICC.2007

[4] Xin C, Xie B, and Shen C. "A novel layered graph model for topology formation and routing in dynamic spectrum access networks"[C]. DySPAN 2005, Baltimore, MD, United States, 2005: 308-317

[5] Cognitive Radio Cognitive Network Simulator http://stuweb.ee.mtu.edu/~ljialian/

[6] 3GPP2, cdma2000 High rate packet data air interface Specification, TS C.S20024 V2.0, October 2000.

[7] O.B. Akan, I.F. Akyildiz, ATL: an adaptive transport Layer for next generation wireless internet, IEEE Journal on Selected Areas in Communications (JSAC)22 (5) (2004) 802–817.

[8] I.F. Akyildiz, Y. Altunbasak, F. Fekri, R. Sivakumar, AdaptNet: adaptive protocol suite for next generation wireless internet, IEEE Communications Magazine 42 (3) (2004) 128–138.

[9] I.F. Akyildiz, X. Wang, W. Wang, Wireless mesh networks: a survey, Computer Networks Journal 47 (4)(2005) 445–487.

[10] I.F. Akyildiz, Y. Li, OCRA: OFDM-based cognitive radio networks, Broadband and Wireless Networking Laboratory Technical Report, March 2006.

[11] L. Berlemann, S. Mangold, B.H.Walke, Policy-based reasoning for spectrum sharing in cognitive radio networks, in: Proc. IEEE DySPAN 2005, November 2005, pp. 1–10.

[12] V. Brik, E. Rozner, S. Banarjee, P. Bahl, DSAP: a protocol for coordinated spectrum access, in: Proc. IEEE DySPAN 2005, November 2005, pp. 611–614.

Venkaiahnaidu.Adapa received B.Tech degree from JNTU Kakinada affiliated college of JNTU Kakinada in 2009 and he is currently pursuing his Masters degree in computer networks and security from K.L.University (2010-2012). His research interests are cognitive wireless mesh networks.

Sk.MansoorRahaman received B.Tech degree from affiliated college of JNTU Hyderabad in 2010 and he is currently pursuing his master's degree in computer networks and security from K.L.University (2010-2012). His research interest are wireless mesh networks and wireless sensor networks.

Addressing Asymmetric Link in Wireless Mesh Networks

Ashok Kumar. S*, Krishnammal. N**

*II M.E CSE, Sri Shakthi Institute Of Engineering and Technology, Anna University,Coimbatore.
**Asst.prof CSE, Sri Shakthi Institute of Engineering and Technology, Coimbatore.

Abstract

In a mesh network, each node acts as a router/repeater for other nodes in the network. These nodes can be fixed pieces of network infrastructure and/or can be the mobile devices themselves. In such networks, because of the heterogeneous transmission range of the clients and routers, link asymmetry problem exists. Link asymmetry poses several challenges such as the unidirectional link problem, the heterogeneous hidden problem and the heterogeneous exposed problem. These challenges degrade the network performance. The proposed approach addresses these challenges and eliminates the unidirectional link in the network layer.

Index Terms—heterogeneous hidden/ exposed problems, link asymmetry, unidirectional link, mesh networks.

I. INTRODUCTION

A wireless mesh network (WMN) is a mesh network created through the connection of wireless access points installed at each network user's locale. Each network user is also a provider, forwarding data to the next node. The networking infrastructure is decentralized and simplified because each node need only transmit as far as the next node. Wireless mesh networks (WMNs) are dynamically self-organized and self-configured, with the nodes in the network automatically establishing an ad hoc network and maintaining the mesh connectivity. WMNs are comprised of two types of nodes: mesh routers and mesh clients. Other than the routing capability for gateway/bridge functions as in a conventional wireless router, a mesh router contains additional routing functions to support mesh networking.

Through multi-hop communications, the same coverage can be achieved by a mesh router with much lower transmission power. To further improve the flexibility of mesh networking, a mesh router is usually equipped with multiple wireless interfaces built on either the same or different wireless access technologies. In spite of all these differences, mesh and conventional wireless routers are usually built based on a similar hardware platform. Mesh routers have minimal mobility and form the backbone for mesh clients. Thus, although mesh clients can also work as a router for mesh networking, the hardware platform and software for them can be much simpler than those for mesh routers.

Mesh networking (also called "multi-hop" networking) is a flexible architecture for moving data efficiently between devices. In a traditional wireless LAN, multiple clients access the network through a direct wireless link to an access point (AP); this is a "single-hop" network. In a multi-hop network, any device with a radio link can serve as a router or AP. If the nearest AP is congested, data is routed to the closest low-traffic node. Data continues to "hop" from one node to the next in this manner, until it reaches its final destination. The transmission range of the mesh router is usually larger than the transmission range of the mesh client. This indicates that link asymmetry exists in the mesh access network.

Link asymmetry causes numerous challenges such as the unidirectional link problem, the heterogeneous hidden problem and the heterogeneous exposed problem which degrade the network performance. In the network layer, an algorithm is developed to establish the local route spanning tree for each mesh client to solve the unidirectional link problem. With the spanning tree, the mesh router and mesh clients can be connected via multihop communication. To address the hidden terminal problem a new control frame delay to send (DTS) is introduced. DTS is used to avoid collision on demand by cancelling the transmission of a heterogeneous hidden terminal. The rest of this paper is organized as follows: In Section II, we present the overview of the problems. In Section III, we present our approach in detail. In Section IV, the link asymmetry problem is addressed. The simulation results are shown in section V. In Section IV, we conclude this paper and outline our future research direction.

II. PROBLEMS OVERVIEW

The transmission range of the mesh router is usually larger than the transmission range of the mesh client. Hence, link asymmetry exists between the mesh router and the mesh client. The link asymmetry raises the following three problems: 1) unidirectional link problem; 2) heterogeneous hidden problem; and 3) heterogeneous exposed problem.

1) Unidirectional link problem

A unidirectional link arises between a pair of nodes in a network when only one of the two nodes can directly communicate with the other node. The clients with small transmission range cannot respond to routers after receiving requests from routers. Consider Fig.1. A unidirectional link exists between the router R and client G. The router R initially sends Request To Send (RTS) signal to client G. Due to the small transmission range of client G, it cannot send Clear To Send (CTS) signal to respond router R. This problem leads to

incorrect topological information and misbehavior of routing protocols, which commonly assume that the links of the network are bidirectional.
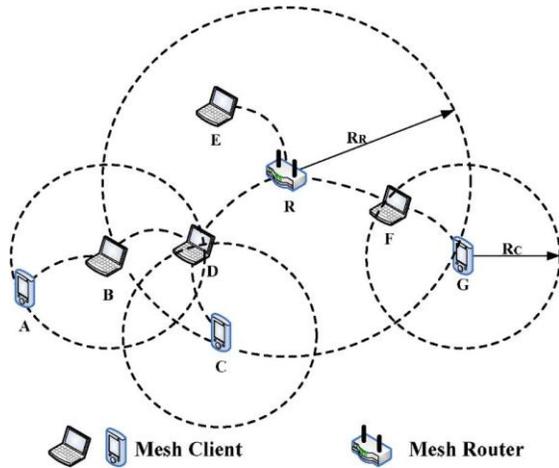

Fig.1 Wireless Mesh Network

2) Heterogeneous hidden problem.

For many wireless technologies in typical ad hoc networking environments, the interference range is larger than the associated coverage range. In such an environment, a node A that does not receive an CTS message from a node B may transmit a packet that will collide with the reception at node B. The reason is that node B may be within the interference range of node A, while node A is outside the transmission range of node B. This problem is called as the hidden part of the interference-range hidden/exposed terminal problem. This refers to collisions raised by accessing the channel from B-node, which cannot be silenced by the G-node CTS frame. Consider Fig. 1. The router R is a heterogeneous hidden terminal of client B. To send data to client B, the router R sends RTS to client B initially. Due to the heterogeneous transmission range the router R cannot receive CTS from client B. Hence, a collision occurs if router R accesses the channel when client B receives data from client A. The heterogeneous hidden terminal increases the possibilities of data collision across nodes and hence the network performance and throughput is affected.

3) Heterogeneous exposed problem.

This refers to the decline of spatial reuse of wireless channel that is caused by clients, which are forced to remain silent by the router's CTS. However, their data transmission will not interfere with the data transmission of the router that sent the CTS. As shown in Fig. 1, both clients C and D are heterogeneous exposed terminals for router R, because clients B and C are within the transmission range of router R and they remain silent by receiving router R's CTS. However, if router R is receiving data from client F or client G, the communication between client B and client C will not affect router R.

III. RELATED WORK

In this section, the proposed approach is described in detail.

1) Basic Handshake and Channel Reservation Operations:

Two new control frames DTS and N-ACK are introduced in our approach. Hence, there are three basic handshake operations in our approach. *RTS/CTS/DATA* is used to handle normal data transmission. When a node needs to send DATA, it first checks the data channel and the control channel. When both channels are idle and the idle time lasts longer than the period of time that is equal to short interframe space, RTS can be transmitted through the control channel. By receiving RTS, if the channel condition allows it to receive DATA, the destination node now replies to RTS by CTS. After receiving CTS from the destination node, the DATA is sent through the data channel. *RTS/DTS/Backoff/ .../retransmit* is used when the channel condition of the destination does not meet the requirements for receiving DATA. After receiving DTS from the destination, the source node will delay its data transmission and retry after backoff. This way, the chance of collision can be largely reduced. The *RTS/CTS/DATA/N −ACK/Backoff/.../Retransmit* is used to provide the reliability for DATA transmission. In case of any collision on the data channel, the destination will send N-ACK to the source. After the backoff procedure, retransmission will recover the collided DATA frame. In addition to these basic handshake operations, channel reservation operations are also performed in this approach. To implement the channel reservation, the network allocation vector (NAV) is used to determine how long the channel will be occupied. Each node maintains three NAVs. In particular, NAVC is used to monitor the control channel. Transmitting a control frame is forbidden when NAVC is positive. NAVS and NAVR are used to manage sending and receiving operations on the data channel. When NAVS is positive, RTS is not allowed to transmit and when NAVR is positive data receiving is forbidden. A Duration field is appended to each control frame to support channel reservation. Based on handshake operations, when handshake is conducted on the control channel, the duration information appended in each control frame can be used to update three NAVs (i.e., NAVC, NAVS, and NAVR) for channel reservation.

IV. ADDRESSING LINK ASYMMETRY

1) Addressing the Unidirectional Link Problem:

To address the unidirectional link it is essential to establish multihop routing reserve paths for mesh routers and mesh clients. Establishing reverse paths between mesh routers and mesh clients is a critical task in this approach. For each mesh client Ci, it is not essential to set up paths connecting it to all routers in the network. Instead, each client Ci only needs to establish paths to connect routers within a range of RR. We define these routers as a set names as R(ci). As shown in Fig. 2, client C1 establishes a path consisting of routers R2, R3, and R4 within the range of RR via clients C4, C2, and C3. We can see that C1 does not establish a path with router R5, because R5 is not within C1's range of RR. Obviously, the local route spanning tree LRST(c1) consists of C1 as its root and routers in R(C1), i.e., (R1, R2, R3, and R4) as its leafs.
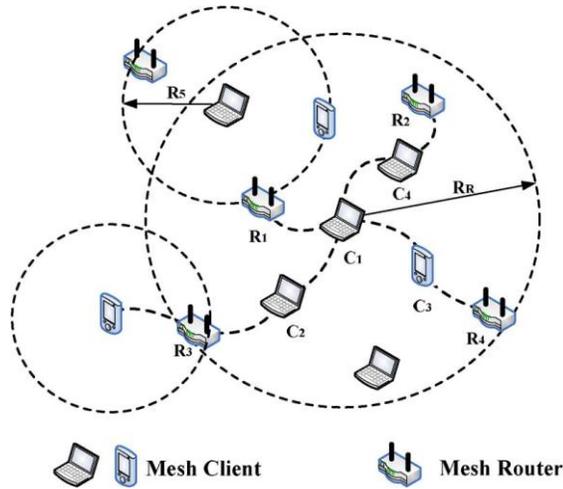
Fig. 2. Example of local route spanning tree.

There are three steps for establishing LRST and addressing the unidirectional link problem. In the first step, a bidirectional table is built for each node to determine whether a unidirectional link exists between a pair of router and client. In the second step, mesh clients are connected to mesh routers by discovering multihop paths. In this step, LRST can be formed in each mesh client. In the third step, the topological information of LRST is used to address the unidirectional link problem. The detailed procedures are presented below.

Step 1: Establishing the bidirectional neighbor table. The bidirectional neighbor table is used to determine whether a unidirectional link exists between the router and the client. The router and the client may periodically need to broadcast Hello packets. In the following, we define $Hello(c_i)$ or $Hello(r_i)$ as the Hello packet sent by client c_i and router r_i, respectively. This way, when a mesh router receives $Hello(c_i)$ or $Hello(r_i)$, it knows that c_i or r_i are its bidirectional neighbors. By receiving $Hello(c_i)$, a mesh client can also ensure that c_i is its bidirectional neighbor.

However, when a $Hello(r_i)$ is received by a mesh client, it cannot determine whether these routers are its bidirectional neighbors or not. To address this problem, $Hello(r_i)$ should append the bidirectional neighbour table of r_i. According to the neighbor list appended in $Hello(r_i)$, a mesh client can determine whether routers are its bidirectional neighbor or not. Hence, the bidirectional neighbor table is formed. In addition, a sequence number is also appended in $Hello(r_i)$. It is used to indicate the freshness of information. Once the router generates a Hello packet, the sequence number will increase. When a $Hello(r_i)$ is received by c_i and r_i is the bidirectional neighbor of this client, a sequence number that is larger than the sequence number maintained for r_i will trigger an update for the record of r_i in the bidirectional neighbour table.

Step 2: Establishing a reverse path to connect the router and the client. When a new bidirectional link between client c_i and router r_i is detected by receiving $Hello(r_i)$ or an update is triggered for r_i due to a fresh sequence number, a connection from c_i to r_i is established. In this case, c_i notifies its

bidirectional neighbors that the path from c_i to r_i has been established by broadcasting a $RSCP(c_i, r_i)$. After receiving $RSCP(c_i, r_i)$, the bidirectional neighbours of c_i will obtain the information that r_i could be connected via c_i. Therefore, those bidirectional neighbors continue to broadcast $RSCP(x, r_i)$, where x belongs to the bidirectional neighbors of c_i. Hence, the bidirectional neighbours of c_i can connect to router r_i. By repeating this process, a distributed LRST will be formed, and routers are connected via multihop paths connected through mesh clients.

Step 3: Eliminating unidirectional link.
The purpose of establishing LRST is to enable control information exchange between the router and the client on unidirectional links for the MAC protocol. By the interactions between the network and link layers in our approach, the MAC protocol can use LRST to route control frames via multihop paths through clients. Fig. 2 shows one simple example. We can observe that the link between router $R3$ and client $C1$ is the unidirectional link. When router $R3$ wants to transmit DATA to client $C1$, it first sends RTS to client $C1$. After receiving RTS from router $R3$, client $C1$ finds that $R3$ is not its bidirectional neighbor according to LRST. Hence, it replies to $R3$'s RTS by CTS via multihop paths through clients. CTS is first delivered to intermediate client $C2$, which then forwards CTS to $R3$. Finally, router $R3$ sends DATA to client $C1$. As we can see, the unidirectional link problem is solved by our proposed mechanisms.

2) Addressing Heterogeneous Hidden Problem:
The main idea to solve the heterogeneous hidden problem is to route control frames, which can either block or delay the router's transmission. One scheme to achieve this goal is to increase the coverage of CTS sent by client . Another scheme is to delay the transmission of router on demand. This scheme is based on the fact that collision may only occur when the heterogeneous hidden terminals access the channel and the client receives DATA. Obviously, the on demanded scheme incurs much less overhead. We adopt the second scheme in our approach.
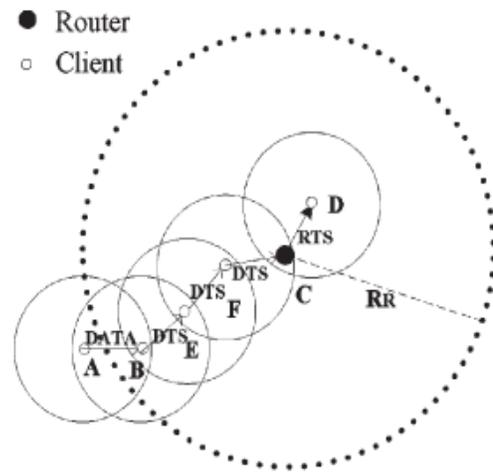

Fig. 3(a). Addressing heterogeneous hidden problem.

In particular, each client senses potential collision by listening to the control channel. If RTS from the router is received by the client on the control channel when it is receiving DATA, the client can ensure that the DATA transmission from that router will collide with the DATA to be received. In this case, DTS is forwarded via a multihop path through mesh clients to the heterogeneous hidden terminal, i.e., the router, to cancel its current data transmission.
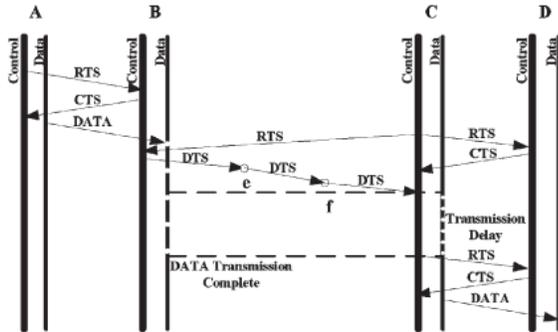

Fig. 3(b). Solution for the heterogeneous hidden problem.

Fig. 3(a) shows a simple example, where client B is receiving DATA from A after RTS/CTS handshake. At this moment, router C tends to send DATA to client D; it sends RTS on the control channel to D, and D replies to CTS. Nevertheless, router C will not immediately transmit DATA on the data channel, but it will wait a period of time to see if there is any DTS that has arrived. When receiving RTS from router C, client B finds that DATA transmission will collide with the DATA to be received. Meanwhile, client B finds that the link between router C is unidirectional according to its LRST; hence, it transmits DTS to client E, and client E forwards the DTS frame. Finally, router C receives DTS, cancels its transmission, and retries according to the Duration field in the DTS frame. Hence, the heterogeneous hidden problem is solved by the proposed mechanisms.

3) Addressing the Heterogeneous Exposed Problem:
Because of the large transmission range of the router, the CTS frame from the router may block data transmission from clients. Hence, the problem becomes how to decrease the coverage of the CTS frame. Our idea is to limit the effective coverage of the CTS frame from the router. Based on the LRST established in the network layer, each client can determine whether the router is its bidirectional neighbor or not. Hence, when the client receives CTS from the router, it can process CTS in different ways than those listed here. When the router is not its bidirectional neighbor, CTS will be ignored. If the router is the bidirectional neighbor, the CTS frame will be processed. Fig. 4(a) shows one simple example. We can see that clients *D* and *E* are heterogeneous exposed terminals when router *B* replies CTS to client *A*.
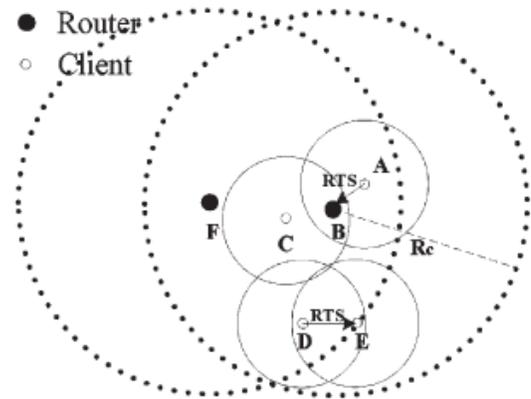
Fig. 4(a). Addressing heterogeneous exposed problem.

Using our mechanism, client *C* and router *F* will be blocked after receiving CTS from router *B*, because they are bidirectional neighbors based on LRST. However, clients *D* and *E* find that router *B* is not their bidirectional neighbor based on LRST. Hence, they simply ignore CTS and initialize RTS for data transmission. Hence, the heterogeneous exposed problem is solved by the proposed mechanisms.
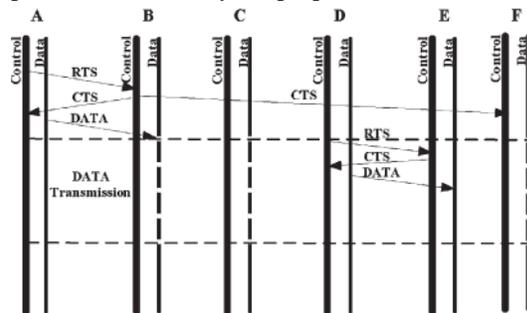
Fig. 4(b). Solution for the heterogeneous exposed problem.

V. PERFORMANCE EVALAUTION
The network capacity and impact of collision are the main aspects considered for the network. To evaluate network capacity, we consider the metric *aggregated one-hop throughput*, which is defined as the total number of packets delivered to the destinations. To measure the impact of collision, we consider the metric *efficiency of data delivery ratio*, which is defined as the ratio of the aggregate one-hop throughput to the number of transmitted packets.

The constant bit rate (CBR) traffic model is used in this simulation as it is a very popular traffic model and has been widely used in the simulation of the MAC protocol. Fig. 5 and 6 shows the simulation results in terms of throughput for CLSM and IEEE 802.11. The throughput of CLSM steadily increases, whereas the throughput of the IEEE 802.11 protocol rapidly decreases when the traffic load increases.
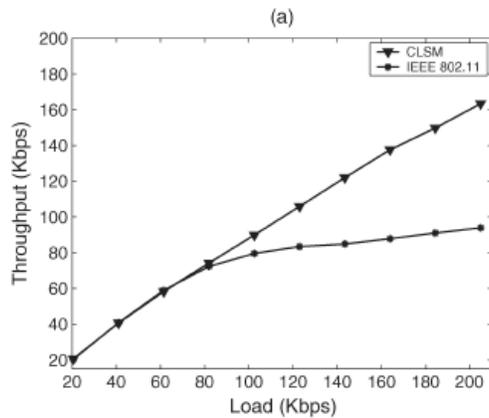
**Fig. 5.** Simulation result: Heterogeneous hidden problem
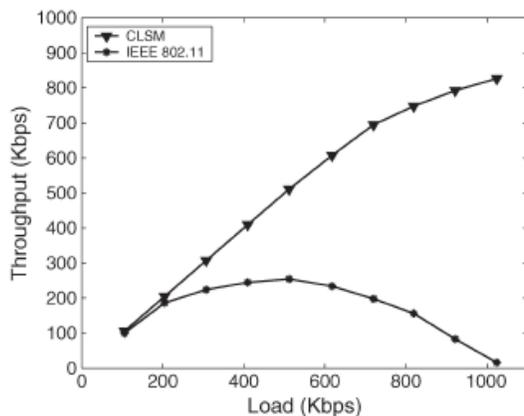
Fig. 6. Simulation result: Heterogeneous exposed problem

VI. CONCLUSION

In this paper, the problems raised by link asymmetry in a wireless mesh network are addressed. The unidirectional link that exists in the network layer is eliminated and the heterogeneous hidden and exposed terminal problems are solved. This approach increases the network performance and throughput which is validated through simulations.

REFERENCES

[1] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Commun. Mag.*, vol. 43, no. 9, pp. S23–S30, Sep. 2005.
[2] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," *Comput. Netw.*, vol. 47, no. 4, pp. 445–487, Mar. 2005.
[3] P. Wang and W. Zhuang, "A collision-free MAC scheme for multimedia wireless mesh backbone," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3577–3589, Jul. 2009.
[4] B. S. Manoj, P. Zhou, and R. R. Rao, "Dynamic adaptation of CSMA/CA MAC protocol for wide area wirelessmesh networks," *Comput. Commun.*, vol. 31, no. 8, pp. 1627–1637, May 2008.
[5] IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ISO/IEC Std. 8802-11: 1999(E), Aug. 1999.
[6] Y. Y. Su, S. F. Hwang, and C. R. Dow, "An efficient cluster-based routing algorithm in ad hoc networks with unidirectional links," *J. Inf. Sci. Eng.*, vol. 24, no. 5, pp. 1409–1428, 2008.
[7] C. H. Yeh, H. Zhou, P. H. Ho, and H. T. Mouftah, "A variable-radius multichannel MAC protocol for high-throughput low-power heterogeneous ad hoc networking," in *Proc. IEEE GLOBECOM*, Dec. 2003, pp. 1284–1289.
[8] C. H. Yeh, "The heterogeneous hidden/exposed terminal problem for power-controlled ad hoc MAC protocols and its solutions," in *Proc. IEEE 59th VTC*, May 2004, pp. 2548–2554.
[9] N. Poojary, S. V. Krishnamurthy, and S. Dao, "Medium access control in a network of ad hoc mobile nodes with heterogeneous power capabilities," in *Proc. IEEE ICC*, Helsinki, Finland, Jun. 2001, pp. 872–877.
[10] V. Shah, E. Gelal, and S. V. Krishnamurthy, "Handling asymmetry in power heterogeneous ad hoc networks," *Comput. Netw.: Int. J. Comput Telecommun. Netw.*, vol. 51, no. 10, pp. 2594–2615, Jul. 2007.
[11] B. J. David and A. M. David, "Dynamic source routing in ad hoc wireless networks," *Mobile Comput.*, vol. 353, no. 4, pp. 153–181, 1996.
[12] H. Zhai and Y. Fang, "Physical carrier sensing and spatial reuse in multirate and multihop wireless ad hoc networks," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006, pp. 1–12.
[13] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Info. Theory*, vol. 46, no. 2, Mar. 2000, pp. 388–404.
[14] X. Du, D. Wu, W. Liu, and Y. Fang, "Multiclass routing and medium access control for heterogeneous mobile ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 270–277, Jan. 2006.

AUTHORS

Mr. S. Ashok Kumar received B.E degree in CSE from Anna University, Chennai and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. His research interest includes Networks.

Mrs. N.Krishnammal received B.E degree in ECE from Anna University, Chennai and received M.E degree under Anna University of Technology, Coimbatore and puruing PHD in Networks under Anna University Coimbatore. She is currently working as a assistant professor in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Her area of interest is Networks.

Optimizing Efficiency of Square Threaded Mechanical Screw Jack by Varying Helix Angle

Tarachand G. Lokhande[1], Ashwin S. Chatpalliwar[2], Amar A. Bhoyar[3]

*(Assistant Professor, Department of Mechanical Engineering, SSPACE, Wardha (M.H) INDIA
** (Associate professor, Department of industrial engineering, SRKNEC, Nagpur (M.H) INDIA.
*** (Assistant Professor, Department of Mechanical Engineering, SSPACE, Wardha (M.H) INDIA

ABSTRACT

This paper deals with Optimization of efficiency of square threaded mechanical screw jack with respect to different helix angle .mathematical model has been done to quantify the effect of varying helix angles. It is concluded that efficiency become large and optimum at helix angle 3.6952936[0] for 10000 Kg screw jack.

Keywords - Optimization, Mechanical screw jack, square thread, helix angle, mathematical modeling.

I. INTRODUCTION

Screw jack is a portable device use to raise or lower the load. The movement of nut on the screw jack is similar to the movement of a weight on an inclined plane, since when one thread is developed, it is an inclined plane as shown in figure 1 and nut taking place of the weight. The base of the inclined plane will be equal to πd_m.

α = Helix angle &   p = pitch threads
Then $\tan \alpha = P/\pi d_m$
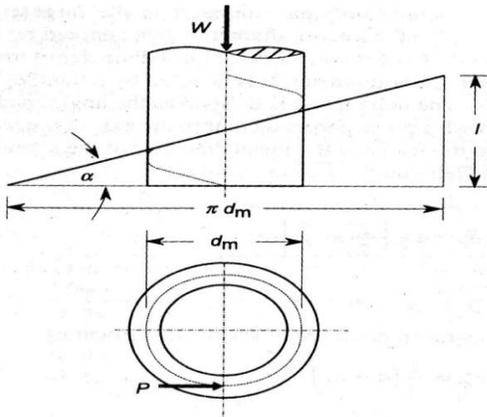
Figure 1: Inclined plane
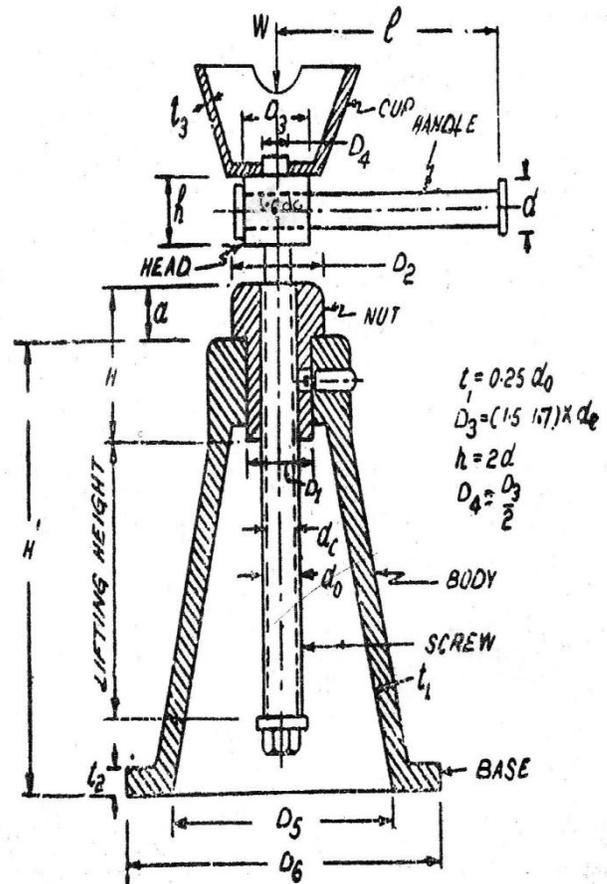
Figure 2: screw jack.

For multi start screws P is replaced by lead 'l'. The square threads are mainly used in screw jacks. Optimization is the act of obtaining the best result under given circumstances. The ultimate goal of optimization is to either minimize the undesired effect or maximize the desired benefit. Optimization can be defined as the process of finding the conditions that gives the maximum or minimum value of a function. The optimum design for mechanical elements should be conducted on the basis of most significant quantity to be minimized or maximized. The optimization is done either to maximize desirable effect or to minimize the undesirable effect [13]. In the earlier study B.R.Singh & Onkar Singh a prototype of air turbine was developed & its functionality was ensuring by testing it [1]. R.K. Jain & B.D. Gupta have conducted study on experimental investigation on the rotary furnaces. They have done modeling & simulation of energy (fuel) consumption of L.O.O. fired rotary furnace. [2] The multilayer feed forward modeling method (with two hidden layers) of artificial neural network continued in MATLAB Software is used for modeling & simulation of energy (fuel) consumption. Claudiu Valantin, Hozumi Goto Hisanori Abiru have optimized screw worm jack by imposing constraints on it. It was optimized for maximum efficiency & minimum overall size [3]. Mastaka Yashimura described a total system optimization method from conceptual design stage to final product realization [4]. N.S. Yalvoi & A.M. kats applied complex approach to the optimal design of machine building production to optimize essentially all structure parameter [5]. Sunil Jha and Manoj Modi have made use of dimensional analysis to investing the effect of electrical & physical parameter on the material removal rate & surface roughness of a die sinking (PMEDM) for hole drilling application. A new mathematical modeling been developed by using dimensional analysis to study the effect of various process parameters like duty cycle, voltage, and powder concentration etc. on material removal rate (MRR) and surface roughness. The model has been validated using the experimental values [6] S. Charles P. Venugopal & R. Bright Reginold Raja conducted tests by varying different cutting parameters and effect of built up base nose size and frank wear analyzed by using design of experiments. Mathematically models were developed to quantify this effect [7]. Hem Chandra Reddy, and Krishna Reddy have given a mathematical flow model which has been developed to study the performance of helical capillary tubes which simulated a situation closer to that existing in practice [8]. B. R. Singh and Omkar Singh optimize the output of vane type air turbine. To different injection angles. Mathematical model than to quantity the effect of varying injection angle, expansion due to isobaric, adiabatic, expansion and steady flow work of high pressure air he concluded that power output becomes large and optimum at injection angle 60^0 to 75^0 and decreases their after [9]. This methodology has been used in the project singh B.R. & onkar singh have carried out parametric evaluation of vane angle on the performance of Novel air turbine.

TABLE 1
PARAMETERS ON WHICH EFFICIENCY DEPENDS

Symbol	Parameter
dc	Core diameter of screw.
do	Outer diameter of screw.
dm=do+dc/2	Mean diameter of screw.
Pb	Bearing pressure.
N	No of threads.
$\Phi = \tan^{-1}\mu$	Friction angle.
$\mu = \tan\Phi$	Coefficient of friction between nut & screaw.
$\alpha = \tan^{-1}(P/\pi dm)$	Helix angle.
$T = W*dm/2*\tan(\alpha+\Phi)$	Torque to be raised or lowered.
ή	Efficiency
P	Pitch of threads

2. MATHEMATICAL MODELLING

The movement of the nut on screw is similar to the movements of a nut on the inclined plane man machine plane as shown in fig & nut taking the place of weight. The base of inclined plane will be replaced by $= \pi d_m$, α- Helix angle & p= pitch of the threads.

$$\text{Tan}\alpha = p/\pi d_m \text{ ---------}1$$

$$d_m = (d_o+d_c)/2 \text{ ---------}2$$

The torque to be transmitted is given by

$$T = w*dm/2*\tan(\alpha+\Phi) \text{ ------------}3$$

Considering the wear of nut,

$$W = \pi/4*(do^2 - dc^2)*P_b*n \text{ ---------}4$$

The efficiency of screw jack to lift the load is given by

$$\text{Efficiency } ή = \tan\alpha /\tan(\alpha+\Phi) \text{ --------}5$$

$$\text{Tan}(\alpha+\Phi) = \tan\alpha/ή \text{ ------------}6$$

Substituting equation (4) & (6) in eq. (3)

$$T = \pi/4*(do^2 - dc^2)*P_b*n*dm/2*\tan\alpha/ή.$$

$$T = \pi/4*(do-dc)*(do+dc)*Pb*n*dm/2*P/\pi dm/n.$$

$$= P*dm*Pb*n*P/4ή$$

$$ή = P*dm*Pb*n*P/4T. \text{ ---------}7$$

$$T = Pb*P^2*dm*n/4ή \text{ ------------}8$$

$$T = \pi/4*(do^2 - dc^2)*P_b*n*dm/2*\tan(\alpha+\Phi).$$

$$= \pi/4*(do^2 - dc^2)*P_b*n*dm/2*\tan\alpha/ή.$$

Tanα = 2Tή/(π*[(dodc)/2]*(do+dc)/2*Pb*n*dm)

Tanα = 2Tή/(π*P/2*dm*Pb*n*P)

Tanα = 4Tή/(π*Pb*dm^2*n*P)

α = tan^{-1}[4Tή/ (π*Pb*dm^2*n*P)--------**9**

TABLE II
FOR 10000 KG LOAD

T (kg)	Friction angleΦ^0	Helix angle α^0	dc	do	Pitch	No of threads	% ή	Critical load (Wcr) kg
4504.0848	11.309932	3.123846	32	38	6	25	26.22	21548.121
4623.81	11.309932	3.6952936	31	38	7	20	27.42	20675.41
5739.6757	11.309932	3.3122712	40	48	8	12	22.07	35797.218
5938.994	11.309932	3.1685559	42	50	8	12	22.30	39584.094
6138.3697	11.309932	3.0367887	44	52	8	11	20.64	43616.904

3. RESULT AND DISCUSSION
Based on the various input parameter listed in table-1 & mathematical model, the effect of helix angles on various parameter studied efficiency, critical load, core diameter , outer diameter, torque to be transmitted, no. of thread & pitch of threads. Friction angle of screw jack is 11.309232^0, coefficient of friction μ=0.20 for whole study & bearing pressure were kept constant throughout the study.

[B] Helix angle vs. no of threads.

Figure 4: Helix angle vs. no. of threads

As the helix angle increases no of threads increases to some extent or helix angle decreases as no of threads decrease.

[A] Helix Angle Vs. Critical load

Figure 3: Helix angle vs. Critical load

As the helix angle increases critical load decreases. The minimum critical load at α is 3.6952936.

[C] Helix angle vs. efficiency.
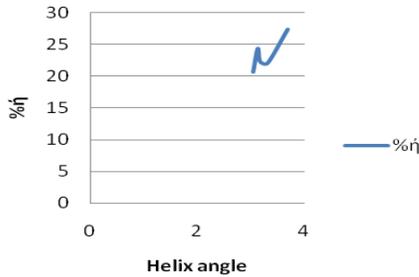
Figure 5: Helix angle vs. efficiency

As the helix angle increases the efficiency increases. But after certain value the efficiency also decreases. The efficiency is maximum for helix angle 3.6952036^0.

[D] Helix angle vs. turning moment.
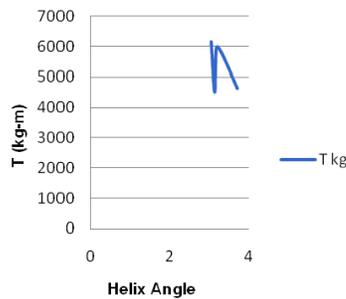
Figure 5: Helix angle vs. Turning moment

As helix angle increases the turning moment also reduces. The turning moment is maximum at 3.0367887^0.

[E] Helix angle vs. outer diameter.
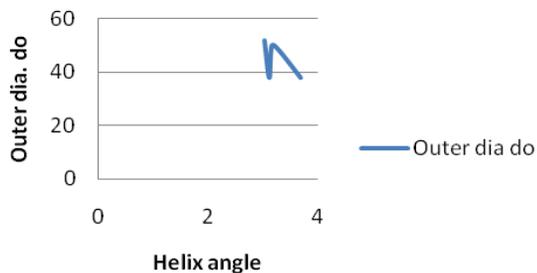
Figure 6: Helix angle vs. outer dia.

As the helix angle increases the outer dia. decreases.

[F] Helix angle vs. core dia. (dc)
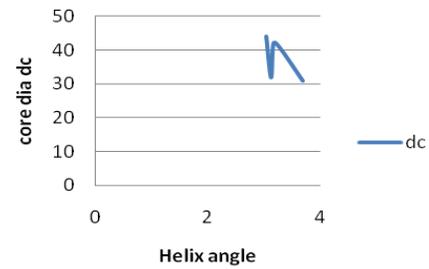
Figure 7: Helix angle vs. Core Dia (dc)

As the helix angle increases core diameter decreases.

[G] Helix angle vs pitch.
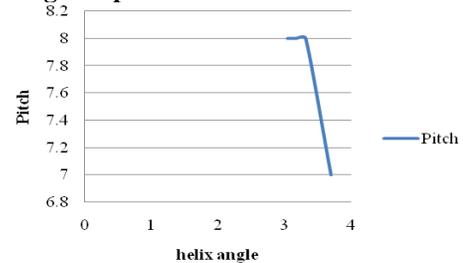
Figure 7: Helix angle vs. Pitch

As the helix angle increase pitch does not change it remains constant. It changes slightly.

CONCLUSION

Based on the input parameter & result obtain the following conclusion are drawn in references with the efficiency of square thread mechanical screw jack for 10000 Kg.

(i)   As the helix angle increases the efficiency increases up to certain limit after which it decreases.

(ii)   As helix angle increases the critical load decreases.

(iii)   As helix angle increases the no of threads decreases.

(iv)   As helix angle increases the turning moment reduces.

(v)   As helix angle increases the outer diameter decreases.

(vi)   As helix angle increases the core diameter decreases.

(vii) As helix angle increases the pitch does not change it remains constant up to certain value & then it reduces.

Thus from above study it is noticed that the jack efficiency becomes large and optimum when helix angle is 3.6952936^0 when the coefficient of friction is 0.20 and bearing pressure between nut & screw is 150 kg/cm^2.

NOMENCLATURE

dc : Core diameter of screw or minor diameter of screw.

do : Outer diameter or major

diameter of screw.

dm = (do+dc)/2 :Mean diameter of the screw.

Pb : Bearing Pressure bet^n nut & screw.

n : Number of threads.

$\Phi = \tan^{-1}\mu$: friction angle.

$\mu = \tan\Phi$: Coefficient of between nut & screw.

$\alpha = \tan^{-1}(P/\pi dm)$.Friction angle.

$T = W*dm/2*\tan(\alpha+\Phi)$: Torque to be applied.

W: Load to be lowered.

ή = Efficiency of screw jack.

P : Pitch of threads.

fyc : Yield stress in compression.

fyt : Yield stress in tension.

fys : Yield stress in Shear.

D1: Diameter of nut ay bottom.

D2: Diameter of nut ay top.

H : Height of nut.

Wcr : critical load in kg.

d : Diameter of lever.

l : length of lever.

fb : Bending stress.

t_1 : Thickness of body at top.

h = 2d

t_2 : Thickness of body at bottom.

t_3 : Thickness of cup.

D5 : Inside diameter of body.

D6 : Outside diameter of body.

Wdes = βW =1.3W: Design load.

REFERENCES

[1] Singh B.R. and Singh Onkar, 2008, " Development of a Vaned type novel Air Turbine", International Journal of Mechanical Engineering Science (The manuscript was received on 21st December 2007 and was accepted after revision for publication on 3rd June 2008), International Journal of IMehE 222 Part C, pp 1419-1426.

[2] R. K. Jain and B. D. Gupta, "Optimization and Simulation of Energy (Fuel) Consumption of L.D.O. Fired Rotary Furnace Using Back Propatation Algorithm", International Journal of Mechanical Engineering, IJME, July – December 2009, 2, pp 177 - 184.

[3] Claudiu Valentin Suciu, Hozumi Goto, Hisanori Abiru, "Modeling and Simulation of a Screw-Worm Gear Mathematical Transmission to Achieve its Optimal Design Under Imposed Constraints", ICPPW, pp. 160-165, 2009 International Conference on Parallel Processing workshops, 2009.

[4] Masataka Yoshimura, "System Design Optimization for Product Manufacturing", Concurrent Engineering 2007; 15; 329.

[5] N.S. Yalovoi and A. M. Kats, "Complex Optimal Design of Centrifugal Pumps", Chemical and Petrochemical Engineering, Vol. 35, Nos. 11-12, 1999.

[6] Sunil Jha and Manoj Modi, "Modeling and Analysis of Powder Mixed Electric Discharge Machining", International Journal of Mechanical Engineering. IJME July – December 2009, 2, pp 219 – 223.

[7] S. Charls, P. Venugopal and R. Bright Reginold Raja, " Analysis and Modeling of Built up Base Nose Size and its Effects on Tool Flank Wear When Turning Steels", International Journal of Mechanical Engineering. IJME July – December 2009, 2, pp 69 – 74.

[8] Hemchandra Reddy, K. & Krishna Reddy, V, "Numerical Studies on Helical Non Adiabatic Capillary Tube", International Journal of Mechanical Engineering. IJME January – June 2009, 2, pp 37 – 41.

[9] Singh B.R. and Singh Onkar, "Optimization of Power Output of a Vaned Type Novel Turbine with Respect to Different Injection Angles", International Journal of Mechanical Engineering. IJME July – December 2009, 2 PP 205-211.

[10] Singh B.R. and Singh Onkar, 2008, "Parametric Evaluation of Vane Angle on Performance of Novel Air Turbine", Journal of Science, Engineering & Management, SITM, December 2008, 2, PP 7-18.

[11] V. B. Bhandari 2006, "Introduction to Machine Design", Tata McGraw-Hill Publishing Company Limited, New Delhi, India, ISBN – 0 – 07 – 043449 – 2.

[12] Joseph E Shigley, Charls R Misckhke, Richar G. Budynas, Kith J. Nisbett, "Mechanical Engineering Design", Tata McGraw-Hill Publishing Company Limited, New Delhi, 2008, ISBN 13: 978-0 – 07 – 066861 – 4, ISBN 10: 0 – 07 – 066861 – 2

[13] Ray C. Johnson, "Optimum Design of Mechanical Elements", A Willey – Interscience Publication John Wiley & Sons Inc. New York, ISBN - 0 – 0471 – 03894 – 6, [pp 1, 2, 3,178 -188, 193 – 196.] .

[14] Shingresu Rao, "Engineering optimization (theory and practice)", New Edge International Publishers Pvt. Ltd. New Delhi 1996.

[15] P.C.Sharma and D. K. Agrawal, "Machine Design", Katson Publishing House, B. D. Kataria & sons, Opp. Clock Tower, Ludhiana, Punjab.

Finite Element Analysis of Single Lap Bolted joint under bending load

Amar A. Bhoyar[1], Sharad S. Chaudhari[2], Tarachand G. Lokhande[3]

*(Department of Mechanical Engineering, Y.C.C.E Nagpur, India)
** (Department of Mechanical Engineering, Y.C.C.E Nagpur, India)
*** (Department of Mechanical Engineering, S.S.P.A.C.E, Wardha, India)

ABSTRACT

Finite element analysis (FEA) is a computational tool that can be used for calculating forces, deformations, stresses and strains throughout a bonded or bolted structure. This paper presents the insight of stress analysis in a single lap bolted joint under bending load. Present work includes finite element approach to study the results of failure of threaded fasteners under bending load. The Experimental data is taken into reference for the analysis purpose. Modeling of single lap joint and then analysis has been performed. Results obtained after analysis was then articulated which show good agreement. A three-dimensional finite element model of a bolted joint has been developed using Pro-E wildfire 4.0 and ANSYS 11 commercial package. Finally, severe areas were identified and confirmed with the stress distribution results from simulation. The load-displacement curves can be compared easily with experimental data. The FEA outputs, such as stress and strain, can be used with failure criteria to predict failure.

Keywords - **Lap Joint, Fastener failure, bending load, Stress Distribution, Modeling, Analysis**

I. INTRODUCTION

A bolted joint is one of the joining techniques employed to hold two or more parts together by the help of nut and bolt to form an assembly in mechanical structures [1]. The single lap joint shown in fig.1 is obtained by overlapping two plates and then fastening with nut-bolt. Bending load is the internal load generated within a bending element whenever a pure moment is reacted or a shear load is transferred by beam action from the point of application to distant point of reaction. Moreover it is an internal tensile or compressive longitudinal stress developed in a beam in response to curvature induced by an external load [12]. Kovács *et al.* [2] carried out an experimental study on the behavior of bolted composite joints. The composite base columns were investigated under cyclic loading. Su and Siu [3] analyzed the nonlinear response of a bolt group under in-plane loading using the numerical method. In order to predict the physical behaviors of the structure with a bolted joint, simulation with three dimensional finite element models is desirable. With the recent increase in computing power, three dimensional finite element modeling of a bolted joint in bending has become feasible. T. N. Chakherlou, M. J. Razavi, A. B. Aghdam [4] carried out the study of variation

of bending force and its concomitant effects on the performance of bolted double lap joints subjected to longitudinal loading. The results unanimously revealed a gradual initial reduction of bending force followed by a significant increase as the longitudinal load was increased. Also affected, was the load transfer mechanism in the joint resulting in variation of friction force between the plates, but in a different trend compared to bending force. Maggi et al. [11] also demonstrated using the same software how variations of geometric characteristics in bolted end plate could change the connection. Investigation on failure of threaded fasteners due to vibration spans nearly sixty years. Sparling [5] found that the fatigue life of the bolt could be amplified appreciably by tapering the nut thread form for the first few engaged threads measured from the loaded face of the nut. It also stated that truncating the threads improved fatigue life but reduced the static load capacity. Nevertheless, experimental studies in the late 1960s by Junker [6] demonstrated that loosening and failure both becomes more rigorous when the joint is subjected to dynamic loads perpendicular to the thread axis (shear loading). The most extensively used apparatus for experimental study of loosening under dynamic shear load is the transverse vibration test apparatus developed by Junker [6]. Bolts, apart from the load of preliminary tension (bending force) and torsion (if hydraulic or electric torque wrench is not applied), are subject to extra stretching or relieving, and Bending [7]. It was recently shown [8] that a fastener could turn loose under dynamic shear loading as a result of accumulation of localized slip in the form of strain at the fastener contacts surfaces. Krishnamurthy [9-10] was among the pioneers to perform FEM analyses of extended and flush end-plate connections, which included elastic material behavior. Krishnamurthy et al. (1979) extended the earlier study to develop a two-dimensional plane stress FEM of bolted connection and conducted some physical tests to confirm the analyses. When a dynamic load is applied as is the case in many assemblies, failure may occur due to fatigue and vibration induced loosening. These failures can be avoided by appropriate joint design where the finite element studies can be supportive.
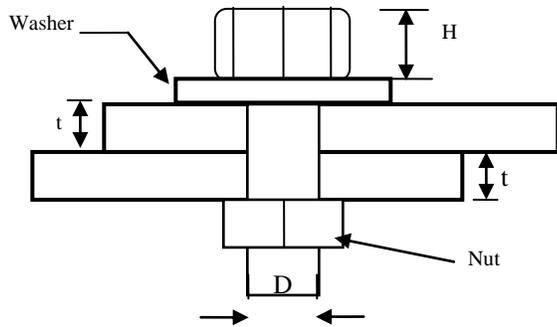
Fig 1. Single lap bolted joint

II. GEOMETRY MODELING

A single lap bolted joint consists of a bolt, a nut, washers and two plates as shown in Fig 1. Most repeatedly bolts used in machining are made to SAE standard J429 (Unified Engineering, 2006). The bolt and nut selected is a hex bolt of SAE Grade 5, 0.5 inches in diameter. Fig 2(a) illustrates the geometry of the standard SAE Grade 5 hex bolt. The mechanical properties and dimensions of the bolts are shown in Table 1. Washers are not used in the design of the specimen because of its influence on the accuracy of torque controlling.

The plate material properties used is based on mild steel which is shown in Table 2. The work for this study was carried out using the finite element method. In this work, Pro-E 4.0 modeling software is used for modeling of bolted joint. Finite element software program named ANSYS 11 was used for analysis of single lap bolted joint. In simulation, the geometry scale factor is determined as 1,000 mm. In general, the unit used in ANSYS modeling is listed in Table 3. The geometry of the model is based on two solid plates (170 × 74 × 9 mm), one support plate (40 ×74 × 9 mm) and an unit consisting of a bolt and a nut. The bolt hole of radius 6.15 mm, is edited in the upper and lower plate. The diameter of the bolt shank is assumed to fully fit in the bolt hole. As illustrated, the position of origin is pointed by the arrow; direction represents the length of the plate, y direction represents the width of the plate, and z direction represents the thickness of the plate. The FE model of the test joint was developed using ANSYS 11, which is general-purpose finite element analysis software. A representative finite element mesh of the model used for the study is shown in Fig. 3(b). It consists of a Nut, which fastens the top plate through a threaded insert. The geometry is cut down to include only the key features of the structure. Since the base is assumed to be stiff, only a small region around the threaded insert is modeled, and the nodes on the external surface of this region are controlled shown in fig 3(a)
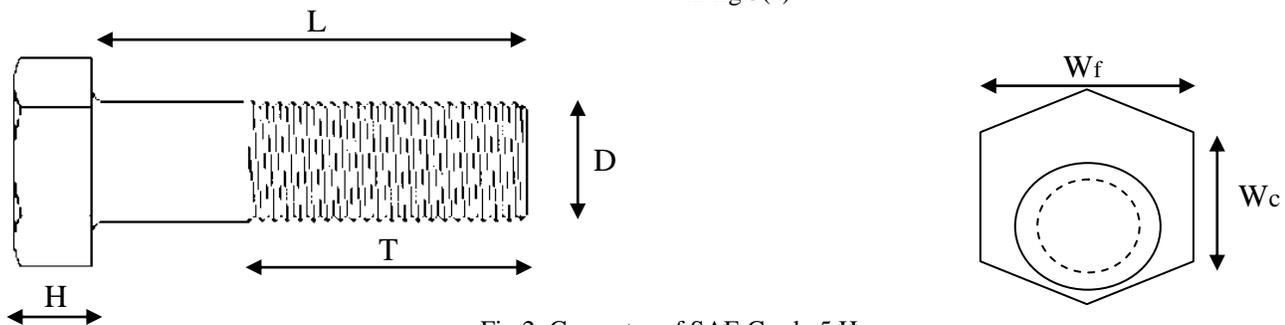
Fig 2. Geometry of SAE Grade 5 Hex

TABLE 1
SPECIFICATIONS OF NUT & BOLTS

Sr.No	Properties	Notation	Dimension/Type
1.	Material	-	Medium carbon steel, quenched and tempered
2.	Supposed length	L	37.8 mm
3.	supposed diameter	D	12.3 mm
4.	Height of bolt	H	7.5 mm
5.	Height of nut	H_n	11.5 mm
6.	Width across flat	W_F	19.1 mm
7.	Width across corners	W_C	22.3 mm
8.	Modulus of elasticity	E	200 Gpa
9.	Poisson's ratio	µ	0.29
10.	Proof strength	-	589 Mpa
11.	least tensile yield strength	-	634 Mpa
12.	least tensile ultimate strength	-	827 Mpa

TABLE 2
SPECIFICATIONS OF THE PLATE

Sr. No	Property	Notation	Type/Dimension
1.	Material	-	Mild steel
2.	Width	w	75 mm
3.	Thickness	t	10 mm
4.	Poisson's ratio	μ	0.29
5.	Modulus of Elasticity	E	200 Gpa

TABLE 3
PARAMETERS AND UNITS USED

Sr.No	Parameter	Unit (SI)
1.	Force	N
2.	Mass	Tonnes
3.	Time	Second
4.	Length	mm
5.	Density	Tonnes/mm
6.	Stress	Mpa

III. MESH GENERATION

The finite element mesh used in this study utilizes a mainly hexahedral mesh, which provides good fallout within a levelheaded runtime. The element used on the bolt and plates is the same with SOLID 92 type Tet10. The total elements in the model are 43,298 and the numbers of nodes are 64,029. Higher order elements are used in the contact regions of the components shown in Fig. 3(c). Finally, the mesh was verified to test the failure of aspect ratio, edge angle, face skew, collapse, normal offset, tangent offset, etc. The verification summary fig.3 (d) showed the number of failures the elements involved. Zero number of failures is desired to minimize error in the analysis process. The position of origin shown in fig.3 (b) is pointed by the arrow; direction represents the length of the plate, y direction represents the width of the plate, and z direction represents the thickness of the plate. These contact regions are overlaid with high order general-purpose contact elements capable of modeling friction. The model includes contact regions between the nut head and the fastened component shown in Fig 3(a), the fastened component hole and the nut body, and between the internal and external threads. Preliminary studies were conducted to conclude the effect of the mesh density on the loosening results. It was found that a relatively coarse mesh utilizing high order elements at the contact regions provided good results in a level headed amount of time. This is not astonishing since loosening results are based on the displacements, which converge with a relatively coarse mesh compared to the mesh required for accurate determination of stresses. Mesh Density of the plates is shown in Fig 3 (b).

BOUNDARY CONDITIONS

The boundary conditions for the single lap bolted joint are decided by selecting nodes as shown in Fig 3(a). In this section, displacement is created to constrain the model from translating when a Bending load is applied. The ideal displacement is applied in a line/curve to uniform the displacement distribution. The locations of displacement are 25 mm and 225 mm in x direction from the origin. However, in this model, the application region is determined by those nodes which are on the desired lines. These nodes are constrained by fixing the degrees of freedom; translations are in x, y, and z directions.
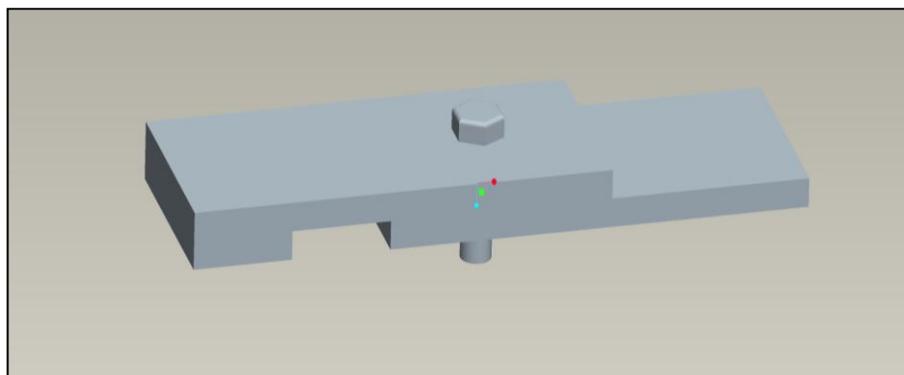

Fig 3(a). Geometry Modeling of Model
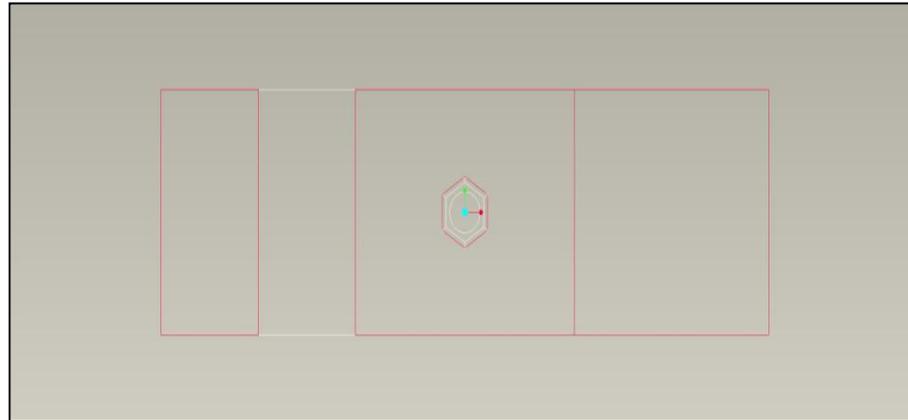
Fig 3(b). Position of Origin
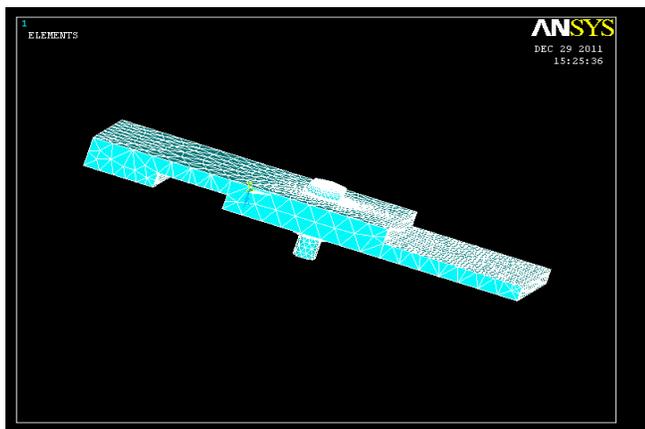
(c)
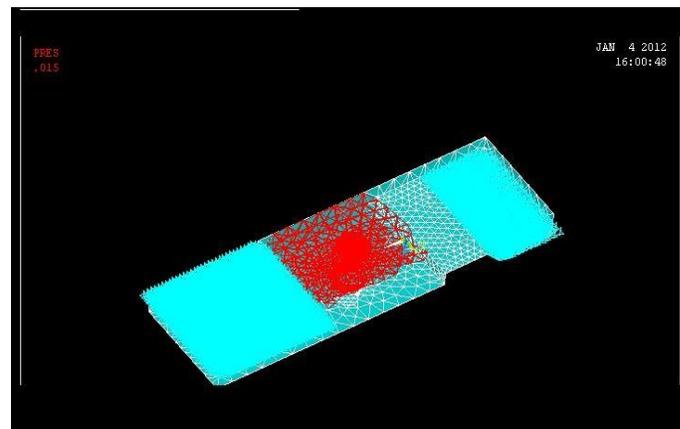
(d)

Fig 3(c) & (d) Meshing of Model & Verification of Mesh

IV. EXPERIMENTAL PROCEDURE [1]

The experimental work was carried out in the Strength Laboratory at UPM to test four-point bending on a single lap bolted joint. The results obtained in the experimental measurements are load-displacement and load-strain data. The experimental load-displacement curves were obtained directly from the testing machine. Strains at selected points on the joint surface were measured while bending forces were controlled manually by the Universal Testing machine. The surface of the lower plate was strain-gauged in the axial direction. Fig 6(b) shows the position of the unidirectional strain gauge, which had a 5 mm gauge length. Axial strain in the plate is observed for every increment of 2 KN. The specimen and jigs were mounted on the universal testing machine as illustrated in Fig 6(c). The bending load was applied by the lower load cell vertically, while the top load cell was fixed. The application of the load was controlled manually to obtain the strains measurement for every increasing load level.

LOADING

In this step, only one type of loads was considered which the bending force is. A bending force of 100 N is modeled by using a negative pressure value. The bending force is applied by uniform pressure onto the overlap area between the bolt nut and the plates. The load locations applied is 85 mm and 165 mm in direction x in origin. The pressure applied on the lap joint can be calculated. The area of contact between two plates is 90 x 74 Sq. mm and the bending force of 100 N is applied. The pressure applied is equal to 0.015015 N/mm^2

V. RESULT AND DISCUSSION

The deformed shape of the finite element model is shown together with the actual deformation in Fig 6(f) & 6(d). The experimental load deflection curves were found to be essentially linear up to and including 8 KN. Since the experimental specimen deformation becomes nonlinear plastic behavior after 8 KN, the shape of the experimental and finite element model with applied load 8 KN is shown for comparison. It can be seen that the finite element model displays similar deformation shape to the experiment. The experimental load-displacement curves were found to be essentially linear until an applied load of 8 KN, so the displacements of the experimental and finite element model were measured over this range. Fig 6(a) shows the calibration curves of load-displacement of the experimental and finite element model under elastic region of three tested samples. Stress distribution of the joint is observed at different load levels to track the stress distribution in each parts of the joint as shown in Fig 6(g). This technique possesses accurate prediction of stress distribution.
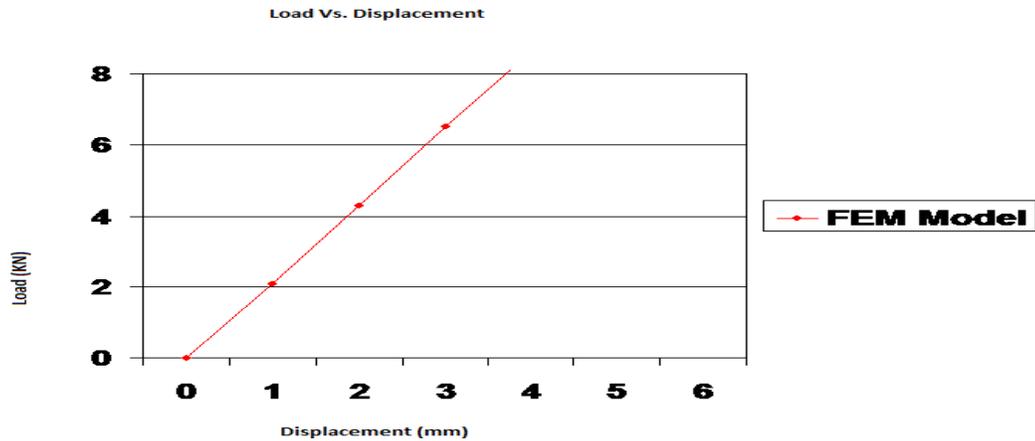
Fig 6 (a). Load Vs. Displacement Graph
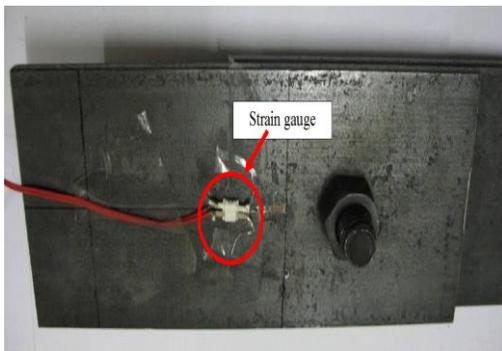
(b)
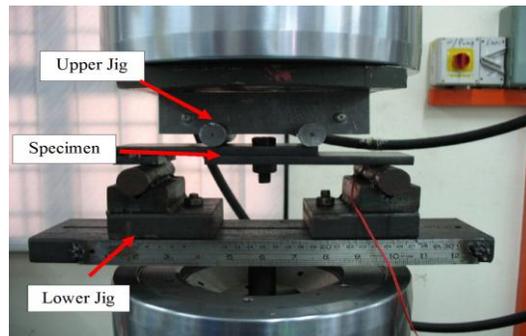
(c)

Fig 6(b) & (c). Strain-gauge set-up & Experimental configuration [1]

The applied load levels used in the analysis are 0 KN, 2 KN, 4 KN, 6 KN and 8 KN. The relative magnitude of the stress distribution on the bolt head decreases with the increasing load level when compared with the stress on the plates. It can be observed that the bolt nut starts to loosen when

Bending force is applied. As expected, higher stress is distributed at the bending and constraint area as illustrated in Fig 6(e). When the specimen experienced 14 KN bending load. Non-symmetrical stress distribution along the x-axis is due to the non-uniform mesh in the model.

(d)
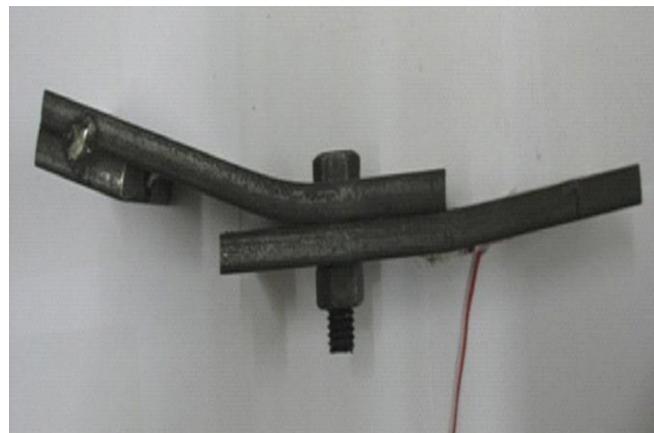
(e)

Fig 6(d) & (e). Experimental specimen & Deformation of Specimen [1]
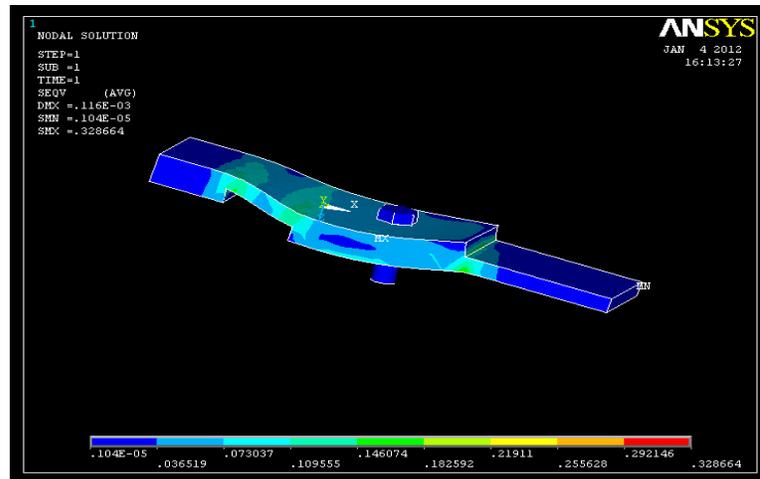
(f)
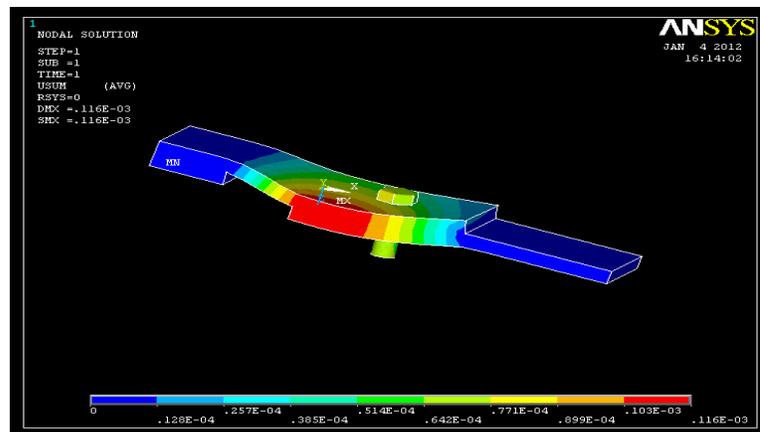
(g)

Fig 6(f) & (g). Progressive Stress distribution & displacement

The trend of stress along the path shows that the maximum stress occurs between the contact of the upper and lower plates. This phenomenon is caused by the combination of contact stress and bending force. This occurs when heavier contact stress between the bolt shank and bolt hole occurs due to bending. Then again, higher stress at the edge of the bolt hole shows agreement with the bending force between the bolt nut and plates. It is interesting to note that the stress changes non-uniformly along the path and increases sharply at the end of the bolt head. This phenomenon is due to the bending force increasing the tension in the bolt head.

Inspite of the critical area occurring on the surface of upper plate of the single lap bolted joint as predicted, it is interesting to investigate the critical area of the bolt and nut, since it was the main component in the joint. From the simulation result, it was observed that the critical areas in terms of Von Mises stress, of the bolt and nut were between the bending surfaces of the bolt head. The plate will experience deformation due to the bolt penetration will load in bending. The joint strength of the bolted joint still relies on the strength of its bolts in both cases when the tension and bending load were applied.

VI. CONCLUSION

A three-dimensional finite element analysis is used to examine the stress distribution of a single lap bolted joint and comparisons were made to the experimental results. These comparisons include bending force and transverse displacement measurement. The result obtained from the simulation analysis shows conformity with experiment analysis and validation of the bolt model was confirmed for an applied load less than 8 KN. A finite element method has

been successfully developed and it can be applied for the prediction of other material, load and size of geometry in four-point bending of a single lap bolted joint. The critical area could be predicted from the simulation analysis and it could save cost of carrying out experimental work. The FEM gives good control of experimental techniques, confirming, complementing and refining the specimen design before commencing experiment tests.

REFERENCES

[1] Aidy Ali*, Ting Wei Yao, Nuraini Abdul Aziz, Muhammad Yunin Hassan and Barkawi Sahari, Simulation and Experimental Work of Single lap bolted joint tested in Bending, *Suranaree J. Science Technology*. Vol. 14 No. 4; October-December 2007

[2] Kovács, N., Calado, L., and Dunai, L. (2004).Behavior of bolted composite joints: experimental study. Construction Steel Research, 60(3-5):725-738.

[3] Ju, S.H., Fan, C.Y., and Wu, G.H. (2004). Three dimensional finite elements of steel bolted connections. Engineering Structures, 26(3):403-413.

[4] T. N. Chakherlou, M. J. Razavi and A. B. Aghdam (2011). On the Variation of Bending Force in Bolted Double Lap Joints Subjected to Longitudinal Loading: A Numerical and Experimental Investigation. *Journal for Experimental Mechanics*, 10.1111/j.1475-13. 05.2010.00795. X

[5] SPARLING, L. G. M., 'Improving the strength of Nut fasteners', *Chart. Mech. Engr,* April 1982, 58-59.

[6] JUNKER GH. New criteria for self-loosening of fasteners under vibration. *SAE Transactions* 1969; 78:314–35.

[7] D. Lehmann, Nonlinar axis force and bending moment distribution in bolt clamped L-flange, Stahlbau 9 (2003) 653-663 (in German).

[8] Gambrell SC. *Why bolts loosen* (Machine Design 1968; 40:163–7)

[9] Krishnamurthy N. Fresh look at bolted end-plate behavior and design. *Engineering Journal, American Institute of Steel Construction* 1978; 15(2):39–49 (second quarter).

[10] Krishnamurthy N, Krishna VR. Behavior of splice–plate connections with multiple bolt rows. Report submitted to the Metal Building Manufacturers Association; February 1981.

[11] Y.I. Maggi, R.M. Conclves, R.T. Leon, L.F.L. Ribeiro, Parametric analysis of steel bolted end plate connections using finite element modeling, J Construct. Steel Res. 61 (2005) 689–708.

[12] V.B.Bhandari, *A Textbook of Design of Machine Elements*, (Tata Mc'Graw hill, 2010-11)

A Cluster Based MARDL Algorithm for Drifting Categorical Data

A Siddhartha Reddy[1], C V V N Varun[1], AnushaAre[2], P.V.V Prasad[2], K. Ruth Ramya[2]

[1](M.Tech student, Department of CSE, K L University, Vaddeswaram, A.P.-522502, India)
[2](Asst.Professor, Department of CSE, K L University, Vaddeswaram, A.P.-522502, India)

ABSTRACT:
Clustering is an important problem in data mining. Most of the earlier work on clustering focused on numeric attributes which have a natural ordering on their attribute values. Recently, clustering data with categorical attributes, whose attribute values do not have a natural ordering, has received some attention. However, previous algorithms do not give a formal description of the clusters they discover and some of them assume that the user post-processes the output of the algorithm to identify the final clusters. Sampling has been recognized as an important technique to improve the efficiency of clustering. However, with sampling applied those points that are not sampled will not have their labels after the normal process the problem of how to allocate those unlabeled data points into proper clusters remains as a challenging issue in the categorical domain. A mechanism named Maximal Resemblance Data Labeling for to kept the every unlabeled data point in to appropriate cluster , the MARDL will exhibits high execution efficiency and it preserves clustering characteristics that is high intra-cluster similarity and low inter-cluster similarity.

Keywords: Data mining, Data labeling, categorical Clustering.

1. INTRODUCTION

Clustering in the computer science world is the classification of data or object into different groups. It can also be referred to as partitioning of a data set into different subsets. Data cluster are created to meet specific requirements that cannot created using any of the categorical levels. One can combine data subjects as a temporary group to get a data cluster. Data clustering is an important technique for exploratory data analysis and has been the focus of substantial research in several domains for decades [8], [9]. The problem of clustering is defined as follows: Given a set of data objects, the problem of clustering is to partition data objects into groups in such a way that objects in the same group are similar while objects in different groups are dissimilar according to the predefined similarity measurement. Therefore, clustering analysis can help us to gain insight into the distribution of data. However, a difficult problem with learning in many real world domains is that the concept of interest may depend on some been explored in the previous works [1,2], hidden context, not given explicitly in the form

of predictive features. In other words, the concepts that we try to learn from those data drift with time [7, 11, and 12]. For example, the buying preferences of customers may change with time, depending on the current day of the week, availability of alternatives, discounting rate, etc. As the concepts behind the data evolve with time, the underlying clusters may also change considerably with time [1]. Performing clustering on the entire time-evolving data not only decreases the quality of clusters but also disregards the expectations of users, which usually require recent clustering results. The problem of clustering time-evolving data in the numerical domain has [3, 4, 5, 6, 10, and 13]. However, this problem has not been widely discussed in the categorical domain with the exception of for Web log transactions. Actually, categorical attributes also prevalently exist in real data with drifting concepts. For example, buying records of customers, Web logs that record the browsing history of users, or Web documents often evolve with
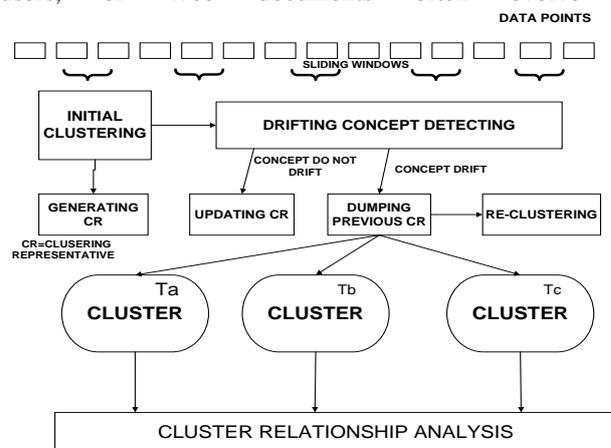
Fig. 1. The framework of performing clustering on the categorical time-evolving data.

 time. Previous works on clustering categorical data focus on doing clustering on the entire data set and do not take the drifting concepts into consideration. Therefore, the problem of clustering time evolving data in the categorical domain remains a challenging issue. As a result, a framework for performing clustering on the categorical time-evolving data is proposed in this paper. Instead of designing a specific clustering algorithm, we propose a generalized clustering

framework that utilizes existing clustering algorithms and detects if there is a drifting concept or not in the incoming data.

The fig shows our entire framework of performing clustering on the categorical time-evolving data. In order to detect the drifting concepts, the sliding window technique is adopted. Sliding conveniently eliminate the outdated records, and the sliding windows technique is utilized in several previous works on clustering time-evolving data in the numerical domain [1,3, 4,12]. Therefore, based on the sliding window technique, we can test the latest data points in the current window if the characteristics of clusters are similar to the last clustering result or not.

1.1. Definition1 (*Node*)

A *node*, dt , is defined as *attribute name + attribute value* .The term *node* which is defined to represent attribute value in this paper avoids the ambiguity which might be caused by identical attribute values. If there are two different attributes with the same attribute value, e.g., the age is in the range *50~59* and the weight is in the range *50~59*, the attribute value *50~59* is confusing when we separate the attribute value from the attribute name. *Nodes [age=50~ 59]* and [*weight=50~59]* avoid this ambiguity. Note that if the attribute name and the attribute value are both the same in the nodes d1 and d2, d1 and d2 are said to be equal. For example, in Figure 3 cluster c1, *[A1=a]* and *[A2=m]* are nodes.

1.2. Node Importance Representative (NIR)

We next describe the novel categorical cluster representative which is named *NIR* (standing for **N**ode **I**mportance **R**epresentative). The basic idea behind NIR is to represent a cluster as the distribution of the nodes, which is defined in Definition1.Moreover, in order to measure the representability of each node in a cluster, the importance of node is evaluated based on following two concepts: (1) The node is important in the cluster when the frequency of the node is high in this cluster. (2) The node is important in the cluster if the node appears prevalently in this cluster rather than in other clusters. The first concept characterizes the importance of the node in the cluster. The rationale for us to adopt the second concept to measure the importance of the node can be explained, where an attribute distribution in the two clusters is given. The node *b* is the most frequent node in the cluster 1. However, in the all data points which contain node *b*, there are only around 40% data points which belong to the cluster 1. In contrast, although the node *c* is less frequent than node *b* in the cluster 1, node c mostly occurs in the cluster 1. Only considering the first concept will cause the importance of node to be high simply because the node is frequent in the database. However, the representability of the node in this cluster is likely to be overestimated because the other clusters also contain this node with high frequency. Consequently, both the two concepts should be employed to evaluate the importance of the node.

Note that the good cluster criteria is high intra-cluster similarity, where the sum of distances between objects in the same cluster is minimized, and low inter-cluster similarity, where the distances between different clusters are maximized. Suppose that there is a node with high frequency in the cluster. This means that most of the data points in the cluster contain this node, and the intra-cluster similarity will be high. Hence, the first concept considers the distribution of the node in the cluster, which can be deemed as the intra-cluster similarity. In addition, suppose that a node occurs in one cluster and does not appear in other clusters. This means that most of the data points which contain this node only occur in this cluster. The distances between different clusters will be large. Hence, the second concept considers the distribution of the node between clusters, which can be deemed as the inter-cluster similarity. Therefore, NIR represents cluster by nodes and the importance of nodes ,which considers both the intra-cluster similarity and the inter-cluster similarity.
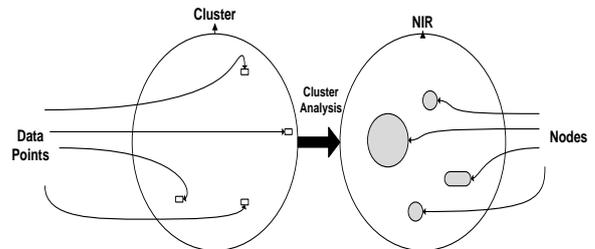

Fig. 2. The concept of NIR to represent a cluster.

As shown in Figure 2, the cluster is represented by NIR.The ellipses in the right side of Figure4 illustrate the nodes in the cluster, and the importance of the nodes is presented by the size of each ellipse. After the process of cluster analysis, a cluster with data points is represented by NIR. To achieve this, the theory of NIR technique is presented below.

Based on the foregoing, cluster ci can be represented by nodes. Each data point in the cluster ci is first decomposed into nodes, and then, the frequency of nodes in the cluster is calculated. The node decomposed from the data point may be equal to the node decomposed from the previous data points. In such cases, the frequency of this node is increased by one. After all the data points are decomposed into nodes in the cluster c_i, suppose that c_i contains t nodes, and each node dk which occurs in the cluster ci is abbreviated by d_{ik} , and, the frequency of node d_{ik} is $|d_{ik}|$. Then, the node importance and NIR are defined as follows.

1.3. Definition 2 (node importance and NIR):

The node importance of the node (d_{ik}) is calculated as the following equations:

$$W(c_i, d_{ik}) = f(d_{ik}) \frac{|d_{ik}|}{\sum_{x=1}^{t} |d_{ix}|} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

$$f(d_{ik}) = 1 - \frac{-1}{\log n} \times \sum_{y=1}^{n} p(d_{yk}) \log(p(d_{yk})),$$

$$\text{Where } p(d_{yk}) = \frac{|d_{yk}|}{\sum_{z=1}^{n} |d_{zk}|} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

The NIR of cluster c_i be represented as a table of the pairs $(d_{ik}, w(c_i, d_{ik}))$ for the all nodes, i.e., $d_{i1} d_{i2}..., d_{it}$, in the cluster c_i.$w(c_i, d_{ik})$ represents the importance of node d_{ik} in cluster c_i with two factors, the probability of d_{ik} in c_i and the probability of d_{ik} in c_i and the weighting function f (d_{ik}). Based on the concepts of the importance of a node, the probability of d_{ik} in c_i is calculated to compute the frequency

of d_{ik} in the cluster c_i, and the weighting function is designed to measure the distribution of the node between clusters based on the information theorem. Entropy is the measurement of information and uncertainty on a random variable. Formally, if X is a random variable, S(X) is the set of values which X can take, and p(x) is the probability function of X, the entropy E(X) is defined as shown in Eq. (3).

$$E(X) = -\sum_{x \in s(x)} P(x) \log (p(x)) \quad\ldots\ldots\ldots\ldots(3)$$

TABLE 1

An example dataset with three clusters and several unlabeled data points.

Cluster c1			Cluster c2		
A1	A2	A3	A1	A2	A3
a	m	c	c	f	a
b	m	b	c	m	a
c	f	c	c	f	a
a	m	a	a	f	b
a	m	c	b	m	a
Cluster c3			**Unlabeled dataset U**		
A1	A2	A3	A1	A2	A3
c	m	c	a	m	c
c	f	b	c	m	a
b	m	b	b	f	b
b	m	c	a	f	c
a	f	a	......	...	....

Explanation: Consider the data set in Table1. Cluster c_1 contains eight nodes ([A_1=a], [A_1=b], [A_1=c],[A_2==m], [A_2==f], etc.).

The node [A_1=a] occurs 3 times (|d1,[A_1=a]| = 3) in c_1 , once in c_2 , and once in c_3 .

The weight of the node [A_1=a], f (d[A_1=a]) = $1 - \frac{-1}{\log 3} (\frac{3}{5}$ $\log \frac{3}{5} + \frac{1}{5} \log \frac{1}{5} + \frac{1}{5} \log \frac{1}{5})$ = 0.135.

The importance of node [A_1=a] in cluster c_1 is: w(c_1 , [A_1 = a]) =0.135∗ $\frac{3}{15}$ = 0.027. Note that in the cluster c1, node [A3=c] also occurs three times. However, this node does not occur in c_2 but occurs twice in c_3.

Therefore, in cluster c_1 , the node [A_3=c] is more significant than node [A_1=a].

Corresponding to the node importance, w (c_1, [A_3 = c]) = f (d [A_3=c]) ∗$\frac{3}{5}$ = 0.387 ∗ $\frac{3}{15}$ = 0.077 > w(c1, [A_1 =a]) = 0.027.

TABLE 2

The NIR table of cluster c1, c2, and c3 in table1

Cluster c_1		Cluster c_2		Cluster c_3	
d_{1j}	$W(d_{1j})$	d_{2j}	$W(d_{2j})$	d_{3j}	$W(d_{3j})$
[A_1=a]	0.027	[A_1=a]	0.009	[A_1=a]	0.009
[A_1=b]	0.004	[A_1=b]	0.004	[A_1=b]	0.007
[A_1=c]	0.005	[A_1=c]	0.016	[A_1=c]	0.011
[A_2=m]	0.009	[A_2=m]	0.005	[A_2=m]	0.007
[A_2=f]	0.005	[A_2=f]	0.016	[A_2=f]	0.011
[A_3=a]	0.014	[A_3=a]	0.056	[A_3=a]	0.014
[A_3=b]	0.004	[A_3=b]	0.004	[A_3=b]	0.007
[A_3=c]	0.077			[A_3=c]	0.052

Although these two nodes both occur three times in cluster c1, node [A_3=c] provides more information on cluster c1 than node [A_1=a] Finally, the *NIR* of cluster c_i can be represented as a table of the pairs (d_{ik}, w(c_i, d_{ik})) for the all nodes in the cluster c_i . The table 2 shows the *NIR* of the three clusters in Table1.

2. RELATED WORK

A survey on clustering techniques can be found in [14] Here, we focus on reviewing the techniques of cluster representative and data labeling on the categorical data, which are most related to our work.

Cluster representative is used to summarize and characterize the clustering result [11]. Since, in the categorical domain, the cluster representative is not well discussed, we review several categorical clustering algorithms and explain the sprite of cluster representative in each algorithm.

In k-modes [15], a cluster is represented by "mode", which is composed by the most frequent attribute value in each attribute domain in this cluster. Suppose that there are q attributes in the dataset. Only q attribute values, each of which is the most frequent attribute value in each attribute, will be selected to represent the cluster. Although this cluster representative is simple, only use one attribute value in each attribute domain to represent a cluster is questionable. For example, suppose that there is a cluster which contains 51% male and 49% female in attribute gender. Only using male to represent this cluster will lose the information from female, which is almost a half in this cluster.

In algorithm ROCK [16], clusters are represented by several representative points. This representative does not provide a summary of cluster, and thus cannot be efficiently used for the post-processing. For example, in the data labeling, the

similarity between unclustered data points and clusters is needed to be measured. It is time consuming to measure the similarity between unclustered data points and each representative point, especially when a large amount of representative points is needed for the better representability.

In algorithm CACTUS [17], clusters are represented by the attribute values. The basic idea behind CACTUS is to calculate the co-occurrence for attribute-value pairs. Then, the cluster is composed of the attribute values with high co-occurrence.

However, this representative does not measure the importance of the attribute values. A cluster is represented only by several attribute values and each attribute value has the equally representability in the cluster. In this paper, we present NIR, which is based on the idea of representing the clusters by the importance of the attribute values because the summarization and characteristic information of a cluster can be obtained by the attribute values. Utilizing the summarization and characteristic information to execute data labeling is more efficient than utilizing the representative points.

Furthermore, data labeling is used to allocate an unlabeled data point into the corresponding appropriate cluster. The technique of data labeling has been studied in CURE [18]. However, CURE is a special numerical clustering algorithm to find non-spherical clusters. A specific data labeling algorithm is defined to assign each unlabeled data point into the cluster which contains the representative point closest to the unlabeled data point. In addition, ROCK [16], a categorical clustering algorithm, also utilizes data labeling to speed up the entire clustering procedure. The data labeling method in ROCK is independent of the proposed clustering algorithm, and is performed as follows. First, a fraction of points is obtained to represent each cluster. Then, each unlabeled data point is assigned to the cluster such that the data point contains the maximum neighbors in the fraction of points from the cluster. Two data points are said to be the neighbor of each other if the Jaccard-coefficient [19] is larger than or equal to the user defined threshold θ. However, the threshold θ in ROCK data labeling is difficult to be determined by users. Moreover, it is time consuming to compute the neighbor relationship between an unclustered data point and all representative points.

2.1. Maximal Resemblance Data Labeling

In this paper a mechanism, named **MA**ximal **R**esemblance **D**ata **L**abeling (abbreviated as MARDL), to allocate each categorical unclustered data point into the corresponding proper cluster. The allocating process is referred to as Data Labeling: to give each unclustered data point a cluster label. The unclustered data points are also called unlabeled data points. Figure 3 shows the entire framework on clustering a very large database based on sampling and MARDL. In particular, MARDL is independent of clustering algorithms, and any categorical clustering algorithm can in fact be utilized in this framework. In MARDL, those unlabeled data points will be allocated into clusters via two phases, namely, the Cluster Analysis phase and the Data Labeling phase. The work doing in each phase is described below.
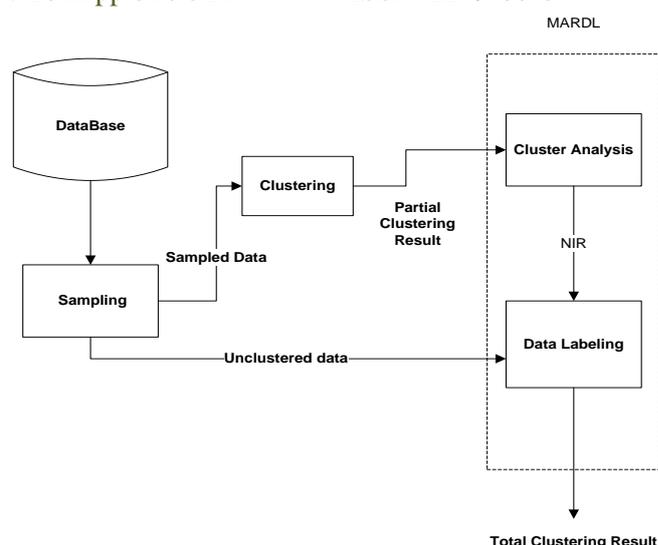
Fig.3. The framework of clustering a categorical very large database with sampling and MARDL.

2.2. Cluster Analysis Phase:

In the cluster analysis phase, a cluster representative is generated to characterize the clustering result. However, in the categorical domain, there Is no common way to decide cluster representative. Hence, a categorical cluster representative, named "**N**ode **I**mportance **R**epresentative" (abbreviated as *NIR*), is devised in this paper.NIR represents clusters by the attribute values, and the importance of an attribute value is measured by the following two concepts: (1) the attribute value is important in the cluster when the frequency of the attribute value is high in this cluster; (2) the attribute value is important in the cluster if the attribute value appears prevalently in this cluster rather than in other clusters. NIR identifies the significant components of the cluster by the important attribute values. Moreover, based on these two concepts to measure the importance of attribute values, NIR considers both the intra-cluster similarity and the inter-cluster similarity to represent the cluster.

2.3. Data Labeling Phase:

In the data labeling phase, each unlabeled data point is given a label of appropriate cluster according to NIR. By referring to the vector-space model [20], the similarity between the unlabeled data point and the cluster is designed analogously to the similarity between the query string and the document. According to this similarity measurement, MARDL allocates each unlabeled data point into the cluster which possesses the maximal resemblance. There are two advantages in MARDL: (1) high efficiency. MARDL is linear with respect to the data size. MARDL is efficient in essence and able to preserve the benefit of sampling on clustering very large database; (2) retaining cluster characteristics. MARDL gives each unlabeled data point a label of the cluster based on the partial clustering result obtained by clustering sampled data set. Since NIR considers the importance of the attribute value, MARDL will preserve clustering characteristic: high intra-cluster similarity and low inter-cluster similarity.

3. MARDL Algorithm

The algorithm MARDL is outlined below, where MARDL can be divided into two phases, the cluster analysis phase and the data labeling phase.

MARDL(C, U):

Clustering result C, unclustered data set U.

Procedure

main ():

The main procedure of MARDL

1. N Table=cluster analysis(C);

2. Data Labeling (N Table, U);

Procedure Cluster Analysis(C): analyze input clustering result and return the NIR hash table

Luster analysis(C)

1. While (C[next]! ='\0') {

2. p [i][j]=C[next];

3. divide Nodes (p [i][j]);

4. Update NF(c[i]);

5.}

6. For (N=d_{i1}; N<=d_{it}; N++)

7. Compute Weight f (d_{ix});

8. For ($C = c_1$ & C<=c_n; C++) {

9. For (N=d_{i1}&N<=d_{it}; N++) {

10. Calculate n (w_i, d_{ix})

11. Add NIR table NTable (d_{ix}, w_i, d_{ix})}}

12. Return NTable;

Procedure Data Labeling (NTable,U): give each unclustered data point a cluster label

13. While (U[next]! ='\0') {

14. U [u][j]=U[next};

15. Divide nodes (p[u][j]);

16. For (N=$c_1 ; N \le c_n; N + +$)

17. Calculate Resemblance(C[m])

18. Give label c[m] to p[u][j] ;}.

The main purpose of the cluster analysis phase is to represent the prior clustering result with NIR. NIR represents cluster by a table which contains all the pairs of a node and its node importance. For better execution efficiency, the technique of hash can be applied on the represented table. Since the node names are never repeated, node is suitable to be a hash key for efficient execution. The main purpose of the data labeling phase is to decide the most appropriate cluster label for each unlabeled data point. Each unlabeled data point is labeled and then classified to the cluster which attains the maximal

resemblance. The resemblance value of the specific cluster is computed efficiently by the sum of each node importance through looking up the NIR hash table q times. After all the resemblance values is computed and recorded, the maximal resemblance value is found, and the unlabeled data point is labeled to the cluster which obtains the maximal resemblance value. Note that after executing the data labeling phase, the labeled data point just obtains a cluster label but is not really added to the cluster. Therefore, NIR table will not be modified in the data labeling phase. This is because the MARDL framework does not cluster data, but rather, presents the original clustering characteristics to the incoming unlabeled data points.

4. CONCLUSION

This paper we formalized the definition of a cluster when the data consists of categorical attributes, and then introduced a fast summarization-based algorithm MARDL. To allocate each unlabeled data point into the appropriate cluster when the sampling technique is utilized to cluster a very large categorical database categorical cluster representative technique, named NIR, to represent clusters which are obtained from the sampled data set by the distribution of the nodes. The evaluation validates our claim that MARDL is of linear time complexity with respect to the data size, and MARDL preserves clustering characteristics, high intra-cluster similarity and low inter-cluster similarity. It is shown that MARDL is significantly more efficient than prior works while attaining results of high quality.

REFERENCES

[1]   C. Aggarwal, J. Han, J. Wang, and P. Yu, "A Framework for Clustering Evolving Data Streams," Proc. 29th Int'l Conf. Very Large Data Bases (VLDB), 2003.

[2]   F. Cao, M. Ester, W. Qian, and A. Zhou, "Density-Based Clustering over an Evolving Data Stream with Noise," Proc. Sixth SIAM Int'l Conf. Data Mining (SDM), 2006.

[3]   D. Chakrabarti, R. Kumar, and A. Tomkins, "Evolutionary Clustering," Proc. ACM SIGKDD '06, pp. 554-560, 2006

[4]   Y. Chi, X.-D. Song, D.-Y. Zhou, K. Hino, and B.L. Tseng,"Evolutionary Spectral Clustering by Incorporating Temporal Smoothness," Proc. ACM SIGKDD '07, pp. 153-162, 2007..

[5]   B.-R. Dai, J.-W. Huang, M.-Y. Yeh and M.-S. Chen, "Adaptive Clustering for Multiple Evolving Streams," IEEE Trans. Knowledge and Data Eng., vol. 18, no. 9, pp. 1166-1180, Sept. 2006.

[6]   M.M. Gaber and P.S. Yu, "Detection and Classification of Changesin Evolving Data Streams," Int'l J. Information Technology and Decision Making, vol. 5, no. 4, pp. 659-670, 2006.

[7]   G. Hulten, L. Spencer, and P. Domingos, "Mining Time-ChangingData Streams," Proc. ACM SIGKDD, 2001.

[8]   A. Jain and R. Dubes, Algorithms for Clustering Data. Prentice Hall,1988

[9]   A.K. Jain, M.N. Murty, and P.J. Flynn, "Data Clustering: A Review," ACM Computing Surveys, 1999.

[10]   O. Nasraoui and C. Rojas, "Robust Clustering for Tracking Noisy Evolving Data Streams," Proc. Sixth SIAM Int'l Conf. Data Mining(SDM), 2006.

[11]   H. Wang, W. Fan, P. Yun, and J. Han, "Mining Concept-Drifting Data Streams Using Ensemble Classifiers," Proc. ACM SIGKDD,2003

.[12]   G. Widmer and M. Kubat, "Learning in the Presence of Concept Drift and Hidden Contexts," Machine Learning, 1996.

[13]   M.-Y. Yeh, B.-R. Dai and M.-S. Chen, "Clustering over Multiple Evolving Streams by Events and Correlations," IEEE Trans. Knowledge and Data Eng., vol. 19, no. 10, pp. 1349-1362, Oct. 2007.in Dynamic Web Sites," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 2, pp. 202-215, Feb. 2008.

[14]   P. Berkhin. Survey of clustering data mining techniques. Technical report, Accrue Software, 2002.

[15]   Z. Huang. Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Mining. Knowl. Discov., 1998.

[16]   S. Guha, R. Rastogi, and K. Shim. ROCK: A Robust Clustering Algorithm for Categorical Attributes. InProc. of the 15th ICDE, 1999.

[17]   V. Ganti, J. Gehrke, and R. Ramakrishnan.CACTUS stering Categorical Data Using Summaries. InProc. of ACM SIGKDD, 1999.

[18]   S. Guha, R. Rastogi, and K. Shim. CURE: An Efficient Clustering Algorithm for Large Databases. InProc. of the ACM SIGMOD Conf., 1998.

[19]   A. Jain and R. Dubes. Algorithms for Clustering Data.Prentiche Hall, 1988.

[20]   R. Baeza-Yates and B. Riberiro-Neto. Modern InformationRetrieval. Addison-Wesley, 1999.

Pradesh, India, in 2009.He is Pursuing M.Tech in Computer Science and Engineering in K.L University, A.P, India during 2010-2012. His research invites Data Mining and Knowledge Discovery

- ANUSHAARE received her B.Tech degree in Computer Science and Engineering from SASTRA University, and M.Tech degree from K.L University .She is currently working as Assistant Professor in K.L University, She is actively engaged in research and publications in the areas of Computer Networks and Data Mining,

AUTHORS

AREGAKUTI SIDDHARTHA REDDY received his B.Tech degree in Computer Science and Engineering from Nalanda Engineering College, Andhra Pradesh, India, in 2010. He is Pursuing M.Tech in Computer Science and Engineering in K.L University, A.P, India during 2010-2012. His research invites Data Mining and Knowledge Discovery

CHAMARTHI VERRA VENKATA NAGA VARUN received his B.Tech in Information&Technnology and Engineering from Nalanda Engineering College, Andhra

EFFICIENT QUERY PROCESSING IN SPARSE DATABASE BY AVOIDING SUBSPACE

S.Dhiyanesh*, Mrs.DeviSelvam**

*II M.E CSE,Sri shakthi Institute Of Engineering and Technology,Anna University,Coimbatore
**Asst.prof CSE,Sri shakthi Institute Of Engineering and Technology,AnnaUniversity,Coimbatore

Abstract

Sparse data are common and available in many real life applications. The Sparse data sets are used in e-commerce application. An E-commerce dataset may have thousands of attributes, but most of the values are null and only a few of while apply to a particular queries to a database. In existing RDBMS objects are conventionally stored using horizontal representation, vertical representation and HoVer representation in this paper. According to the dimension correlation of sparse datasets, a novel mechanism has been developed to conduct query for sparse datasets by improving the 'Hover' technique. Therefore the original data objects are represented in a database format in respective subspaces.

Index Terms: Sparse database, query processing, correlation, subspace, Hover.

I.   INTRODUCTION

With continuous advances in the network and storage technology, there is dramatic growth in the amount of very high-dimensional sparse data from a variety of new application domains, such as bioinformatics, time series, and perhaps, most importantly e-commerce [1], [3], which pose significant challenges to RDBMSs.
The main characteristics of these sparse data sets may be summarized as follows:

High dimensionality: The dimensionality of feature vectors may be very high, i.e., the number of possible attributes for all objects is huge. For example, in some e-commerce applications, each participant may declare their own idiosyncratic attributes for the products, which results in data sets that have thousands of attributes [1].

Sparsity: Each object may have only a small subset of attributes, which is called active dimensions, i.e., significant values appear only in few active dimensions;
In addition, different objects may have different active dimensions. For example, an e-commerce data set may have thousands of attributes, but most of which are null and only a few of which apply to a particular product.

Correlation: Since each object may have only few active dimensions, more likely, similar objects share same or similar active dimensions. For example, in recommendation systems, it is important to find homogeneous groups of users with similar ratings in subsets of the attributes. Therefore, it is possible to find certain subspaces shared by similar objects. In existing RDBMSs, objects are conventionally stored using a horizontal format called the horizontal representation in this paper.

II.   RELATED WORK

In RDBMSs, sparse data sets are typically represented by the horizontal representation. The horizontal representation is straightforward and can be easily implemented; however, it may suffer from schema evolution, column number limitation, and poor storage and query performance incurred by sparsity [2]. Besides the horizontal representation, there are two approaches which can be used to store and conduct query for sparse data sets in an unmodified RDBMS, i.e., the vertical representation [1] and the decomposition storage model [4].

In [1], an alternative of the horizontal representation, i.e., the vertical representation, was investigated. The vertical representation has been used to represent sparse data sets for

Querying [1] as well as representing dense data sets for data mining [5]. In the vertical representation, a single row in the horizontal representation is split into multiple rows. Each row contains an object identifier, an attribute name, and a value.

In addition, the Resource Description Framework (RDF) data of the Semantic Web is commonly stored using this format.

The vertical representation can scale to thousands of attributes, avoid storage of null values, and support evolving schemas. However, writing queries over this representation is cumbersome and error-prone. An inspection of a single row in the horizontal representation becomes a multiway self-join over the vertical table. Another problem of the vertical representation is that it is difficult to support multiple data types. One approach is to create a separate table for each data type, but the approach makes the query rewriting more complicated.

There are some other works about vertical partition of horizontal tables where the subspaces are generated according to the characteristics of the fixed query workload over the data set. On the contrary, our work explores inherent properties of sparse data sets, and our solution is data-oriented, i.e., subspaces are generated according to the distribution of data.

III.   THE HOVER REPRESENTATION

A.   THE HoVer REPRESENTATION

The pure horizontal or vertical representation is introduced that may yield unsatisfactory performance in sparse databases. Therefore, we propose a new representation called HoVer, which can effectively exploit the characteristics of sparse data sets, such as sparsity and dimension correlation. We aim at achieving good space and time performance for storing and querying high-dimensional sparse data sets.

Although the dimensionality of sparse data sets could be very high, up to thousands, a single data object typically has only a few active dimensions, and similar objects have a better chance to share similar active dimensions. A closer inspection of many e-commerce sparse data sets shows that typical e-commerce data sets have a wide variety of items which can be organized into categories and the categories themselves are hierarchically grouped; items that belong to a common category are likely to have common attributes, while those within the same subcategory are likely to have more common attributes. The RDF data [6] also shows that the attributes of similar subjects tend to be defined together. This motivates us to find certain subspaces which are shared by similar data groups, and to split the full space into some lower-dimensional subspaces.

On the other hand, our purpose is to split the full space into subspaces which can yield superior performance for the storage and query of sparse data. These approaches are not suitable for this scenario.

In this section, present sparse data is introduced, these sets using the novel HoVer representation. First, we design an efficient and effective approach to find correlated dimensions. After that, the original full space into subspaces and store the original sparse data set using multiple tables where each table corresponds to a certain subspace.

B.   CORELATED DEGREE REPRESENTATION

Before subspace selection, we first consider how to measure the correlation between two dimensions

Definition 1 (Correlation Table):   The correlation table represents the correlation of dimensions in a sparse data set, which is a super triangle matrix.

Definition 2 (Correlated Degree): The correlated degree is used to measure the correlation between two dimensions in which dimension i(j) is active, and tuples in which dimensions i and j are active simultaneously. According to set theory, it characterizes the number of tuples in which at least one of the two dimensions i and j is active.

C.   SUBSPACE SELECTION

An optimal subspace partitioning should enjoy two properties, i.e., all dimensions are highly correlated intersubspaces while being highly unrelated intersubspaces. If the number of subspaces determined by the user is smaller, dimensions which are not highly correlated may be clustered into the same subspace; hence, the subspace tables are still very sparse.

According to our above analysis, the number of subspaces should be determined by the subspace selection algorithm according to the dimension correlations of the sparse data set. Because the underlying storage and query processing details of the RDBMS may have some influence on the performance [3], there may not exist perfect subspace clustering typically.

Our subspace selection problem can be mapped to the Minimum Clique Partition problem. While mapping each dimension in the sparse data set to a node in the graph, and if the correlated degree between two dimensions is no less than the correlated degree threshold.

Add an edge to link the two corresponding nodes in the graph, and our subspace selection problem is exactly same as the Minimum Clique Partition problem. Unfortunately, the Minimum Clique Partition problem is NP-complete [7], which means that heuristic algorithm is used to approximate optimal partitions which tries to group together correlated dimensions. With a smaller threshold, fewer subspaces will be generated, but the nonnull density of each subspace will be smaller. Actually, the optimal correlated degree threshold varies for different data sets.

D.   SCHEMA EVALUATION

When a new column is added, a new subspace which only contains the new column will be created, and the correlation table should also be updated accordingly. Since the correlation table is incrementally maintained, the new column may be merged to a subspace when subspaces are reorganized. When a column is deleted, then there is a need to delete the column from the corresponding subspace and update the correlation table accordingly.

IV.   QUERY PROCESSING IN HOVER

A.   QUERY REWRITING

Our ultimate purpose is to define horizontally represented views over the HoVer representation. Users typically issue traditional SQL queries over the horizontal view, which can be rewritten into queries over the underlying HoVer representation.

In our work, the dimensions in the original sparse data space are clustered into subspaces, and a horizontal vertically partitioned into subspace tables. In many real-life applications, the dimensions with a high correlated degree are likely to characterize similar topics and have high probability of being accessed together [8]; hence, they should be stored in the same subspace table. Then take advantage of this characteristic and access as few subspace tables as possible in query evaluation. A query rewriting table which records the relationships between the dimensions and the subspaces is essential for query rewriting.

B. EFFECT OF CORELATED DEGREE THROSHOLD

The correlated degree threshold has great influence on the query performance based on the HoVer representation. Using a large threshold, the average size of the subspace tables is small, and the average nonnull density of the subspace tables is large.

In this case, queries over the horizontal representation have high probability of being rewritten into queries over the join of many small subspace tables and the cost of join operations could be high, because the accessed columns have high probability of being distributed into many subspace tables.

In our work, the correlated degree threshold is a tuning parameter and the correlated degree is only characterized by the distribution of the data. Previous work [8], [6] and our experimental study over real-life data sets in Section 5 verify that the correlated degree can well characterize the correlations between dimensions in most of the cases. Besides the subspace partitioning, the diversity of query workload can also significantly impact the overall query performance, as different queries may access various data objects and attributes. Integrating the characteristics of the query workload for subspace selection is a promising alternative to optimize the performance of database for query processing.

V.   CONCLUSION

In this paper, the problem of efficient query processing over sparse databases is addressed. To alleviate the suffering from sparsity and high-dimensionality of sparse data, a new approach is introduced as named HoVer. According to the characteristics of sparse data sets, then vertically partition the high-dimensional sparse data into multiple lower-dimensional subspaces, and all the dimensions in each subspace are highly correlated, respectively. The proposed scheme can find correlated subspaces effectively, and yield superior storage and query performance for conducting queries in sparse databases.

REFERNCES
[1]. R. Agrawal, A. Somani, and Y. Xu, "Storage and Querying of E-Commerce Data," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 149-158, 2001.

[2]. R. Agrawal, R. Srikant, and Y. Xu, "Database Technologies for Electronic Commerce," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 1055-1058, 2002.

[3]. J.L. Beckmann, A. Halverson, R. Krishnamurthy, and J.F. Naughton, "Extending RDBMSs to Support Sparse Datasets Using an Interpreted Attribute Storage Format," Proc. Int'l Conf. Data Eng. (ICDE), p. 58, 2006.

[4]. G.P. Copeland and S. Khoshafian, "A Decomposition Storage Model," Proc. ACM SIGMOD, pp. 268-279, 1985.

[5]. S. Sarawagi, S. Thomas, and R. Agrawal, "Integrating Association Rule Mining with Relational Database Systems: Alternatives and Implications," Data Mining and Knowledge Discovery, vol. 4, nos. 2/3, pp. 89-125, 2000.

[6]. K. Wilkinson, C. Sayers, H.A. Kuno, and D. Reynolds, "Efficient RDF Storage and Retrieval in Jena2," Proc. Int'l Workshop Semantic Web and Databases (SWDB), pp. 131-150, 2003.

[7]. Paz and S. Moran, "Non Deterministic Polynomial Optimization Problems and Their Approximations," Theoretical Computer Science, vol. 15, pp. 251-277, 1981.

[8]. J. Beckham, R. Krishnamurthy, and J.F. Naughton, "The Tradeoff between Horizontal and Vertical Representations of Sparse Data Sets," technical report, http://www.cs.wisc.edu/~sekar/ application/sekar_ecommerce.pdf, 2003.

AUTHORS

Mr. S.Dhiyanesh received B.E degree in CSE from Anna University, Chennai and Currently pursuing M.E degree in Computer Science and Engineering in Sri Shakthi Institute of Engineering and Technology, under Anna University of Technology, Coimbatore. His research interest includes Data Mining.

Mrs. DeviSelvam received the M.E. and B.E. degree in CSE from Avinashilingam University, Coimbatore. She is currently working as Assistant Professor in Department of CSE in Sri Shakthi Institute of Engineering and Technology, Coimbatore. Previously she got a good experience in SRM University, Chennai. She has presented papers in National and International Conferences. She is member of IEEE. Her main research interests include Computer Networks, Mobile computing and Data mining.

Smart Learning through Intelligent Response System

Robin Tommy
Learning and Development
Tata Consultancy Services, Trivandrum, India

Ullas Ravi
Learning and Development
Tata Consultancy Services, Trivandrum, India

Abstract

Smart Learning currently means learning animated lessons brought through the visual media. We present an idea of presenting smart learning on web platform in a wiki model. The content is made more appropriate and understandable to the user using the currently available means of learning mechanisms like graphical simulations, videos and animations. We propose a system where the user understands and learns the concepts of the searched content based on a dynamic animated simulation and other data (wiki, videos and graphical simulations) available in the internet. The user also will be provided an interactive simulation environment to create more awareness on the searched content. The entire dynamic simulation is based on the wiki and other relevant content of the searched data and the interactive environment is simulated from the other contents of the same data available in the internet. The system will have a predefined set of graphics (clip arts and images), emotions and trained patterns defined for a predefined set of content. Here we try to integrate cognitive, affective and psychomotor level of learning and redefine the way SMART learning has evolved. The dynamic simulation and animation is dependent on the target group. In short the proposed system generates a dynamically animated story board from the wiki static content for providing the meaning and conceptual understanding of the searched content followed by a dynamic interactive environment to learn the concept in depth. Along with these the user will also be provided with other relevant content (videos and graphical simulations) from the internet for the searched data. So we propose a smart system for learning which can enable humans to get more depth in their understanding and they interact with the system during the learning process through the intelligent interactive learning mechanism.

Keywords: dynamic animation, simulation, cognitive intelligence, smart system, interactive knowledge.

Introduction

A lot of research has been put into e-learning by many organization and we have a seen an explosion brought about in the internet. In most scenarios the learning is predefined and the user has no engagement or decision in the learning process. Most of the content available online is static and distributed, it is the pain of the user to gather, unite and draw out a conclusion for relevant data knowledge.

In this paper we propose a system which provides a dynamic content evaluation and smart reconstruction of the available content. The content is made available to the user in the friendliest way using dynamic animations depending on his taste and flavor of learning. The content changes for age groups and cognitive level. The system collects all the data available in the open internet space, evaluates the content, construct the most appropriate animated learning model with the available set of graphics and also gives the user with other knowledge resources available in the open forum. As soon as the user completes the cognitive mode of understanding he will be taken into an interactive dynamic environment for creating more awareness and establishing a psychomotor level of understanding.

The currently available systems or e-learning's have static content just thrown to the target group based on the system designer perspective after understanding the requirement of the user. These data cannot be changed easily. The materials inside these systems are priori determined by the designer/tutor. In our system the data is online and it enriches the experience of the user and making him learning more and interact with the data environment.

Early research was based on content based adaptive presentation [1][2]which adaptively presents the content of the page. The contents of the pages are used as clues to derive important features of learners such as their interests, knowledge state etc. learning items are all pre-stored and not changeable; what keeps changing is the order in which course items are delivered, as also described in [3]. Bollacker et al. [4] refine CiteSeer, through an automatic personalized paper tracking module which retrieves each user's interests from well-maintained heterogeneous user profiles. Woodruff et al. [5] discuss an enhanced digital book with a spreading-activation mechanism to make customized recommendations for readers with different types of background and knowledge. McNee et al. [6] investigate the adoption of collaborative filtering techniques to

recommend papers for researchers. They do not address the issue of how to recommend a research paper; but rather, how to recommend additional references for a target research paper. From this perspective, this work is different from our proposed system where learning happens dynamically in interactive environments with intelligent response.

Proposed System Architecture

The following Figure 1 explains the system architecture. When user queries the system the available data on the open internet will be put together. The content will be fetched from various resources giving preference to the ranking and usage of the content. The content can be retrieved from wiki, citreex and other content management platforms. The system draws patterns from the content and extracts the relevant data based on the predictive algorithm. The resulting data will be given to online dynamic interactive system for dynamic animated content generation. With the available cliparts, graphics, videos and audios (from the internet) and emotions the systems starts a learning platform for the user. If the user feels the content is not enough he can regroup his sub-options available and generate a more defined content. After the user completes the learning he will provided with an interactive responsive environment to get an in-depth understanding of the learning
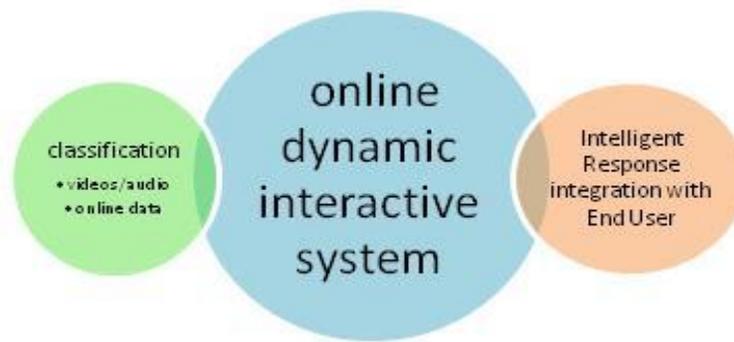
Fig1: Proposed System Architecture

A Running Example

Consider there are two learners Jany and Smitha as shown in figure 2. Both of them belong to different cluster of learner. Jany is observed to browse through the internet for relevant data on Webservices content. She is a novice learner without much background in the topic. She is looking for some contents in the research papers and wiki. Therefore Jany belongs to a group *novice, research, webservice.* Smitha also looking for the same knowledge artifact but is an advanced learner. She is interested in more technical aspects and its implementation aspects. Therefore she can be grouped into the *advanced, technical, webservices.* Here we can see the variant approach in the data needed for the same learning artifact. So static content for both of them is no any value addition to their learning.

The proposed system searches for simple and understandable data for Jany and converts the data into an animated e-learning followed by interactive response learning for giving an insight into the concept. The concepts will be dealt in detail for Jane in his level and provides her an in-depth knowledge taking her to the next level. The same system searches data for Smitha in a more broad sense picking up researched content from IEEE, Citreex and other platforms with applications. The same platform generates an animated e-learning for Smitha giving her more application and research specific data followed by an interactive response system to complete the learning experience.

It is clear from the above example for the same interest, there is a different learning experience generated. Here we propose an individualized system which gives a more precise and relevant data for the target group.
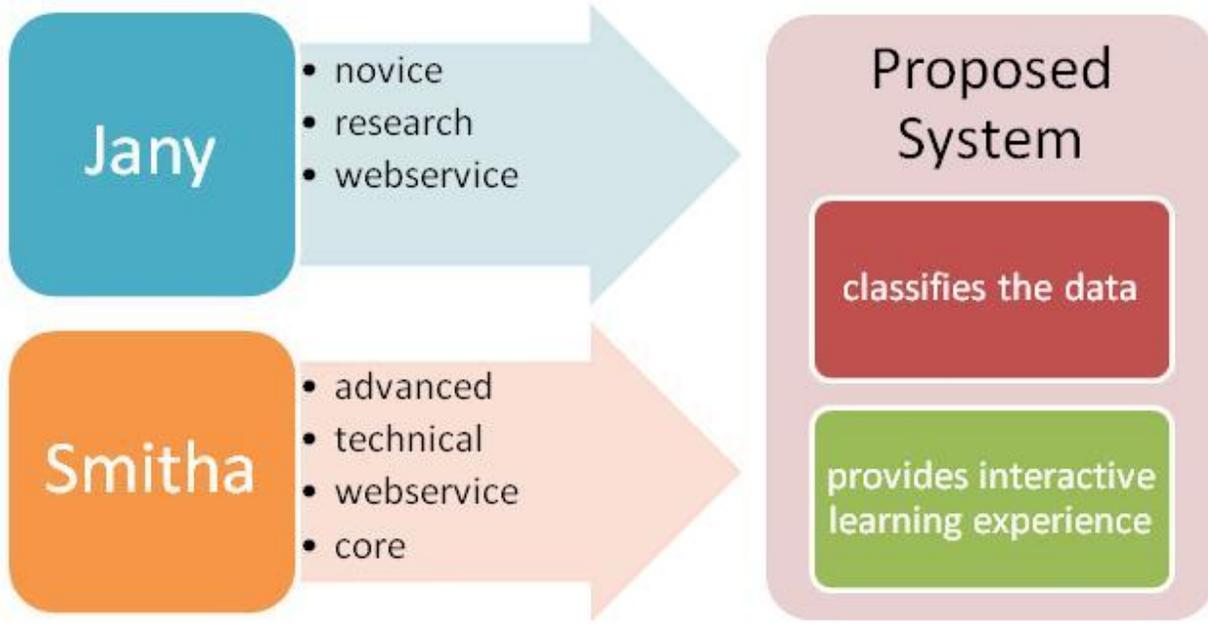
Fig 2: System Example

Conclusion

Current learning systems focus on the system and designer perspectives. The learner is not given the ultimate priority in any scenario. Here we proposed an intelligent system which predicts the right data and provide the exact information for cognitive understanding of the learner in most preferable manner. Here we have made collaboration between the system and the open web giving more possibilities of learning with interactive response environment. It is evident that more research has to be carried out in making this system more learners friendly. It is evident that more rigorous collaborative research should be carried out between researchers from artificial intelligence in education, adaptive hypertext and hypermedia, web information retrieval, data mining, collaborative filtering, user modeling, intelligent user interfaces, computer supported collaborative work etc, in order to achieve these goals.

References

[1] Brusilovsky, P. Adaptive hypermedia. (2001) User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2): 87-110. 2001.

[2] De Rosis, F., De Carolis, B and Pizzutilo, S. (1993) User tailored hypermedia explanations. INTERCHI'93 Conference Proceedings: Conference on Human Factors in Computing Systems, INTERACT'93 and CHI'93, Amsterdam, The Netherlands. 169-170. 1993.

[3] Boyle, C., and Encarnacion, A.O. (1994) MetaDoc: an adaptive hypertext reading system. User Models and User Adapted Interaction. 4, 1-19. 1994.

[4] Bollacker, K.D., Lawrence, S. and Giles, C.L. (1999). A system for automatic personalized tracking of scientific literature on the web. In Proc. ACM Conference on Digital Libraries (DL 1999), 105-113.

[5] Woodruff, A., Gossweiler, R., Pitkow, J., Chi, E. and Card, S.K. (2000) Enhancing a digital book with a reading recommender. In Proc. ACM CHI 2000.153-160. 2000.

[6] McNee, S.M., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S.K., Rashid, A.M., Konstan, J.A and Riedl, J. (2002) On the recommending of citations for research papers. In Proceedings of ACM International Conference on Computer Supported Collaborative Work (CSCW'02), 116-125. 2002.

Sybil Attack In High Throughput Multicast Routing In Wireless Mesh Network

G. Mona Jacqueline[1] and Mrs. Priya Ponnusamy[2]

[1](II M.E., Computer Science And Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore)
[2](Assistant Professor, Department of Computer Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore)

ABSTRACT

Multicast routing is one in which a source sends data to multiple receivers. Multicast routing uses many types of protocols. In this paper ODMRP protocol is used by slightly modifying it. Instead of using hop count for selecting the route, link quality is used so that high throughput is achieved. Also focus is made on attacks that disrupt routing process by modifying the link value, which are very effective against multicast protocols based on high throughput metrics. Other types of attack such as traffic analysis attack, Sybil attack, eavesdropping, selfish attack etc are not focused. In this paper focus is made on Sybil attack. Sybil attack is defined as an attack, in which a malicious node illegitimately takes on multiple identities by spoofing a legitimate node. Detection of Sybil attack which uses multiple identities can be done by nodes that passively monitor the traffic in the network. One method to detect Sybil attack is Passive Ad hoc Sybil Identity Detection (PASID), where a single node can detect Sybil attacks by recording identities, namely MAC or IP address of other nodes. Overtime the node builds a profile which helps in detecting the Sybil attack.

KEYWORDS:- *Wireless Mesh Network, Multicast Routing, Metric Manipulation Attack, ODMRP Protocol, Sybil attack.*

1. INTRODUCTION:

A wireless mesh network (WMN) is a network that consists of radio nodes that are organized in a mesh topology. A mesh network is highly reliable network as it has multiple routes between the sender and the receiver. When an intermediate node becomes inoperable due to mobility or due to some other reasons, the rest of the nodes can still communicate with the help of other intermediate nodes by creating an alternate route.

Multicast routing is one where data is delivered from a source to multiple destinations. There are various protocols for multicast routing [2], [3] and these protocols were proposed for mobile ad hoc networks (MANETS). The protocols mainly focused on hop counts as routing metric for selecting the route between the source and the receiver. The route having least hop count will be selected. But sometimes this leads to selection of a route having poor quality. Recently protocols were developed based upon a metric, the quality of the link [4], [5]. This is referred as link quality or high throughput metrics.

In high- throughput multicast protocols, the nodes at certain time interval will send probes to its neighbors to measure the link quality. If a node wants to send data, a route discovery process is done. The node estimates the link quality by adding its own cost to its adjacent node's cost. The route with a high cost i.e., best link quality is selected. Also, this method is efficient only when there are no attackers. The attackers may modify the original cost so that the route including the attacker node is selected.

The protocol chosen for Wireless Mesh Network for multicast routing is On Demand Multicast Routing Protocol(ODMRP), which chooses the path based upon the hop count. Various types of attacks against the mesh are mesh structure attack, metric manipulation attack etc., In addition to this attack, Sybil attack is made focus in this paper. Sybil attack is an attack where an illegitimate node takes the identity of a legitimate node by spoofing the address of the legitimate node. Moreover, the attacker creates a group of nodes. Following are the main concepts focused in this paper:

- Metric manipulation attack causes severe impact on multicast routing and is of two types: local metric manipulation (LMM), global metric manipulation (GMM) [1]
- A technique called RateGuard[1] is proposed to eliminate the attacker. It combines the measurement-based detection and accusation based reaction techniques.
- Sybil attack is detected using the method called Passive Ad hoc Sybil Identity Detection (PASID).

2. HIGH THROUGHPUT MULTICAST ROUTING

Multihop wireless network is considered where each node participates in data forwarding. Wireless mesh network creates a mesh network by connecting the adjacent nodes. Path selection is done based upon the quality of the link in order to maximize throughput.

2.1 High Throughput Metrics

Routing protocols have used hop count as a path selection metric. The focus has shifted toward high-throughput metrics to maximize throughput by selecting paths based on the quality of wireless links. The quality of the link is found by periodic probing. The total link quality of the path is found by aggregating the link quality of all the nodes in that path.

Some of multicast protocol metrics are SPP, which an adaptation of a unicast metric ETX.

ETX metric: ETX metric [6] was proposed for unicast and it estimates the expected number of transmissions needed to successfully deliver a unicast packet over a link, including retransmissions. ETX is calculated for two nodes A and B by using the following formula:

$$ETX = \frac{1}{d_f * d_r}$$

d_f is the probability taken to deliver data in the forward direction and d_r is the probability taken in the reverse direction. The total value is calculated by aggregating all the ETX.

SPP metric: In multicast both directions are not considered as in unicast. Only the forward direction is considered.

$$SPP_i = d_f$$

3. MESH- BASED MULTICAST ROUTING

Here the ODMRP, ODMRP- HT and how to incorporate the link quality metrics in ODMRP- HT are discussed.

3.1 ODMRP overview:

ODMRP is an on-demand multicast routing protocol for multihop wireless networks, which uses a mesh of nodes for each multicast group. Nodes are added to the mesh through a route selection

3.1.2 JOIN QUERY message

The source node periodically re- creates the mesh by flooding a JOIN QUERY message in the network in order to refresh the membership information and update the routes. JOIN QUERY messages are flooded using a basic flood suppression mechanism, which means nodes only process the first received copy of a flooded message. The node receiving the JOIN QUERY message will again forward the message until the message reaches the destination.
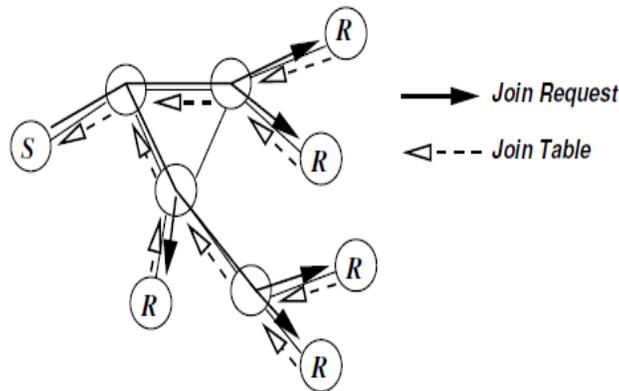
Fig 1: On-Demand Procedure for Membership Setup

3.1.2 JOIN REPLY/ JOIN TABLE message:

When a receiver node gets a JOIN QUERY message, it activates the path from itself to the source node by constructing and broadcasting a JOIN REPLY message that contains entries for each multicast group it wants to join. Each entry has a next hop field filled with the corresponding upstream node. When an intermediate node receives a JOIN REPLY message, it checks if the next hop field of any of the entries in the message matches its own identifier. If so, it makes itself a node part of the mesh (the FORWARDING GROUP) and creates and broadcasts a new JOIN REPLY built upon the matched entries.

Once the JOIN REPLY messages reach the source, the multicast receivers become connected to the source through a mesh of nodes (the FORWARDING GROUP) which ensures the delivery of multicast data. While a node is in the FORWARDING GROUP, it rebroadcasts any non-duplicate multicast data packets that it receives.

3.1.3 Forwarding Group

The forwarding group is a set of nodes that forms a mesh and is capable of forwarding multicast packets. It supports shortest paths between any source and the receiver. All nodes inside the multicast members and forwarding group nodes as shown in fig 2, forward multicast data packets. The multicast receiver can also be a forwarding group node if it is on the path between a multicast source and another receiver as shown in figure 2. The mesh provides richer connectivity among multicast members compared to trees. Flooding redundancy among forwarding group helps overcome node displacements and channel fading. Hence, unlike trees, frequent reconfigurations are not required.

3.2 Incorporating link-quality metrics in ODMRP

To incorporate the new link-quality metrics into ODMRP, the following modification should be done to ODMRP. Each node maintains a NEIGHBOR TABLE that has the costs of the links from its neighbors to itself. The costs are basically a probability value and has a maximum value of one and are periodically updated. In modified ODMRP each node looks up the NEIGHBOR TABLE for the cost of the link from which it received the JOIN QUERY and using this link cost, it updates the cost in the JOIN QUERY packet before rebroadcasting it.

When the JOIN QUERY reaches a receiver, it contains the total cost of the path travelled. The receiver waits for a certain period of time, δ seconds instead of sending JOIN REPLY message immediately on receiving JOIN QUERY message. During this period, it accumulates many duplicate JOIN QUERY packets makes use of the best among them, which is based upon the cost that is the link quality. The path having the highest link quality is selected.

After the expiry of δ seconds, the receiver constructs the JOIN TABLE/ JOIN REPLY message using the stored JOIN QUERY, i.e., the best among all JOIN QUERY packets received during the δ period, and broadcasts the JOIN REPLY to its neighbors

To get high throughput, the forwarding nodes are allowed to forward the duplicate packets. This forwarding is limited to two restrictions. First, a duplicate query is forwarded only if the cost of the path it has traveled is less than that of the minimum cost query received till then. Second, each node sets a timer for a period of $\alpha < \delta$ seconds when it receives the first JOIN QUERY with a particular sequence number. Each node forwards duplicate queries only until the timer of α seconds expires. Choose α as a very small value will lead to minimal path diversity, and a very high value may lead to a high query processing overhead.

3.4 ODMRP- HT:

Fig 2: ODMRP-HT Example having two senders and seven receivers. The value on the link is the SPP metric.

In ODMRP- HT, there will be a route cost field, which determines the link quality. This value is probability value having the maximum value one. ODMRP- HT is similar to ODMRP except, ODMRP- HT selects the route based upon the quality of the link and not on hop count. This selection ensures high throughput. When the source receives the JOIN REPLY message, the source will select the route having the high link value. This method is known as weighted flood suppression. The source will also have the alternate routes. If a certain link fails, the source will choose an alternate route.

4. ATTACKS AGAINST HIGH-THROUGHPUT MULTICAST

The attacker when he attack the mesh, can disrupt the multicast data delivery by either consuming the network resource (resource consumption attacks), by causing incorrect mesh establishment (mesh structure attacks), or by dropping packets (data forwarding attacks). The attacker node on the data delivery path simply drops data packets instead of forwarding them. The attacks are described below:

4.1 Mesh Structure Attacks

Mesh structure attacks disrupt the correct establishment of the mesh structure in order to disrupt the data delivery paths. These attacks can be done by malicious manipulation of the JOIN QUERY and JOIN REPLY messages such as sending large number of invalid JOIN QUERY and JOIN REPLY messages.

For the JOIN QUERY messages, the attacker can spoof the source node and inject an invalid JOIN QUERY message, which causes the paths to be built towards the attacker node instead of the correct source node.

For JOIN REPLY messages, the attacker can drop JOIN REPLY messages to cause its downstream nodes to get detached from the multicast mesh. The attacker can also forward JOIN REPLY to an incorrect next hop node to cause an incorrect path being built.

4.2 Resource Consumption Attack

With ODMRP- HT, the JOIN QUERY message is flooded by the nodes to its neighbours. The attacker node will create a JOIN QUERY message or will simply forward with high frequency which will flood the network. Also, the attacker node will inject more number of JOIN QUERY message to flood the network. Addressing such an attack requires admission control mechanisms which can limit the admission and duration of such groups.

4.3 Metric Manipulation Attack
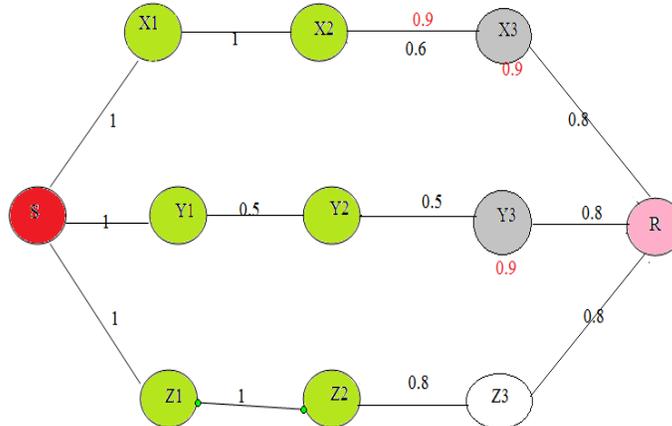
Fig. 3 Example for metric manipulation attack

High throughput multicast routing chooses the route based upon the quality of the link. With metric manipulation attack, the attacker will modify the original link value and will broadcast a false value. There are two types of metric manipulation attack: Local Metric Manipulation(LMM) and Global Metric Manipulation(GMM).

LMM attack: In LMM attack, the attacker will increase the quality of the adjacent link and will broadcast so that the route will be selected along the attacker node. In the above fig, X1 is an attacker. X1 will falsely broadcast the link cost as 0.9 instead of 0.6 so that the route along X1 is selected.

GMM attack: In GMM attack, the attacker instead of broadcasting the adjacent nodes value, will modify the overall accumulated metric and will forward.

5. ATTACK DETECTION AND ELIMINATION

5.1 Measurement based detection

The attack detection is a method that relies on the ability of honest nodes to detect the discrepancy between the expected PDR (ePDR) and the perceived PDR (pPDR). A node can estimate the ePDR of a route from the value of the metric for that route; the node can determine the pPDR for a route by measuring the rate at which it receives data packets from its upstream on that route

If ePDR - pPDR for a route becomes larger than a detection threshold, then nodes suspect that the route is under attack because the route failed to deliver data at a rate with its claimed quality.

5.2 Accusation Based Reaction

A controlled accusation mechanism in which a node, on detecting malicious behavior, temporarily accuses the suspected node by flooding in the network an ACCUSATION message containing its own identity (the accuser node) and the identity of the accused node, as well as the duration of the accusation. As long as the accusation is valid, metrics advertised by an accused node will be ignored and the node will not be selected as part of the FORWARDING GROUP.

5.3 Attack Detection

Attack is detected if there is any variation between ePDR and pPDR.if ePDR- pPDR > δ, then there is an attack. Let n be the number of packets sent by the source and m be the number of packets received by a node in the same period of time. The Wilson estimate requires that n>5, and the confidence interval obtained for pPDR is $(\tilde{p} - e, \tilde{p} + e)$ where

$$\tilde{p} = \frac{m+2}{n+4} \text{ and e} = z \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}}$$

5.4 Attack Reaction

To isolate attackers, protocol uses a controlled accusation mechanism which consists of three components, staggered reaction time-out, accusation message propagation and handling, and recovery message propagation and handling. A react timer is created to act on the adversial node. Based upon this timer the attacker node is accused for a certain period of time.

6. SYBIL ATTACK

Sybil attack is an attack where a node claims multiple identities by spoofing a legitimate node address or ID. In Sybil attack a malicious node behaves as if it were in large number of nodes by impersonating other nodes or simply by claiming false identities.

A Sybil attack is one in which an attacker subverts the reputation system network by creating a large number of pseudonymous entities, using them to gain a disproportionately large influence. A reputation system's vulnerability to a Sybil attack depends on how cheaply identities can be generated, the degree to which the reputation system accepts inputs from entities that do not have a chain of trust linking them to a trusted entity, and whether the reputation system treats all entities identically.

6.1 Prevention of Sybil Attack

Validation techniques can be used to prevent Sybil attacks and dismiss masquerading hostile entities. A local entity may accept a remote identity based on a central authority which ensures a one-to-one correspondence between an identity and an entity and may even provide a reverse lookup. An identity may be validated either directly or indirectly. In direct validation the local entity queries the central authority to validate the remote identities. In indirect validation the local entity relies on already accepted identities which in turn vouch for the validity of the remote identity in question.

6.2 Detecting Sybil Attack

There are several forms of the Sybil attacks are the distributed storage, routing data aggregation, voting, resource allocation and misbehavior detection. Then present the proposed detection technique for these Sybil attacks. The detection has two versions namely a single observer case and a multi-observer case.

6.2.1   Single Node Observer

The single node observer does not require any active probing of suspected Sybil nodes. It operates effectively on a single node that records the identities of nodes. Nodes that are about to detect will record the identities of all other nodes. The time period required for this recording will be depend on the protocol used. After recording the identities, the following steps will be followed by the node.

1. Calculate A_{ij}, the affinity between nodes i and j, as

$$A_{ij} = \left(T_{ij} - 2L_{ij}\right)\frac{T_{ij} + 2L_{ij}}{N}$$

where Tij is the number of intervals in which nodes i and j were observed together, Lij is the number of intervals in which either i or j were observed alone, and N is total number of intervals in the observation period.

2. After calculating the affinity, construct a graph in which the vertices as the identities and the unidirected edges are weighted with the affinity values between them. Only edges that are greater than a specific threshold parameter are included. Using our measure of affinity, we recorded our results using a threshold of 0.1.

3. Depth-first search (DFS) is then run over each vertex to discover the connected components. Each of the components found represents a possible Sybil attacker..

6.2.2 Multiple Node Observers

A subset of the legitimate nodes in the network can share observations periodically using the normal data transmission capabilities of the ad hoc network, and that these nodes can trust each other to perform this task honestly. Each node again tracks all other nodes that it hears over many time buckets. At the end of the observation period, it exchanges the information of what identities were heard during what time periods with the other nodes it trusts in the calculations

7. CONCLUSION

ODMRP- HT, the various types of attacks using high throughput metrics in multicast protocols in wireless mesh networks was considered. Various types of attacks such as metric manipulation attacks, Sybil attack were considered.. Also the Sybil attack was considered. This attack can be detected either using single node or multiple node observer. Further, other types of attacks can be considered. As an initial work, ODMRP-HT protocol is developed.

REFERENCES:

[1] Jing Dong, Reza Curtmola, Cristina Nita-Rotaru "Secure high Throughput Multicast Routing in Wireless Mesh Network", *IEEE Transactions On Mobile Computing, Vol. 10, No. 5, May 2011*

[2] Y.B. Ko and N.H. Vaidya, "Flooding-Based Geocasting Protocols for Mobile Ad Hoc Networks," *Mobile Networks and Applications, vol. 7, no. 6, pp. 471-480, 2002.*

[3] R. Chandra, V. Ramasubramanian, and K. Birman, "Anonymous Gossip: Improving Multicast Reliability in Mobile Ad-Hoc Networks," *Proc. 21st IEEE Int'l Conf. Distributed Computing Systems (ICDCS '01), 2001.*

[4] S.J. Lee, W. Su, and M. Gerla, "On-Demand Multicast Routing Protocol in Multihop Wireless Mobile Networks," *Mobile Networks and Applications, vol. 7, no. 6, pp. 441- 453, 2002.*

[5] D.S.J.D. Couto, D. Aguayo, J.C. Bicket, and R. Morris, "A High-Throughput Path Metric for Multi-Hop Wireless Routing,"   *Proc.ACM MobiCom, 2003.*

[6] S. Roy, D. Koutsonikolas, S. Das, and C. Hu, "High- Throughput Multicast Routing Metrics in Wireless Mesh Networks," *Proc. 26th IEEE Int'l Conf. Distributed Computing Systems (ICDCS),2006.*

[7] Brian Neil Levine, Clay Shields, N. Boris Margolin, "A Survey of Solutions to the Sybil Attack", *Dept. of   Computer Science, Univ. of Massachusetts, Amherst, Dept. of Computer Science, Georgetown University*

[8] John R. Douceur, "The Sybil Attack", *Microsoft Research.*

Application of Genetic algorithm for the determination of optimum machining parameters in turning Al-SiC Metal Matrix Composites

Gururaja Udupa[1], S.Shrikantha rao [2], K.V.Gangadharan[3]

Department of Mechanical Engineering, National Institute of Technology, Karnataka,Surathakal

ABSTRACT

In machining of parts, surface quality is one of the most important requirements. Finish turning using Cubic Boron Nitride (CBN) tools allows manufacturers to simplify their processes by achieving the desired surface roughness.   There are various machining parameters having an effect on the surface roughness, but those effects have not been adequately quantified. Optimum selection of cutting conditions importantly contributes to the increase of productivity and the reduction of costs. This paper attention is paid to this problem in this contribution. A Genetic Algorithm (GA) based approach to complex optimization of cutting parameters is proposed. It describes the multi-objective technique of optimization of cutting conditions by means of genetic algorithm taking into consideration.

Keywords— **Metal matrix composites, Green cutting, Genetic algorithm, Composite machining.**

I. INTRODUCTION

Surface roughness has received serious attention of manufacturers for many years. It has been an important design feature and quality measure in many situations, such as parts subjected to fatigue loads, precision fits, fastener holes and esthetic requirements. In addition surface roughness provide adequate tolerances,which imposes one of the most critical constraints for cutting parameter selection in process planning. While the previous research focused on tolerance study [1-2], this one attempts to develop empirical models with some data mining techniques, such as regression analysis (RA) and computational neural networks (CNN), to help the selection of cutting parameters and the improvement of surface roughness[3-4]. A considerable amount of studies have investigated the general effects of the speed, feed, and depth of cut, nose radius and others on the surface roughness. Empirical models have been developed based on metal cutting experiments using Taguchi designs, and it will include the feed, spindle speed, and depth of cut with different coolants as input variables[5].

The past modeling methods on surface roughness prediction can be classified into two categories: geometric modeling [6] and regression analysis [7]. Geometric modeling is based on the motion geometry of a metal cutting process, regardless the cutting dynamics. Analytical models tend to be general and computationally straightforward. The major drawback of this method is,

they miss other parameters in cutting dynamics including speed, depth of cut and the work piece material in their models[6].On the other hand regression method is a kind of empirical modeling method, is that these studies did not apply the factorial experimentation approach to design the experiments. Therefore, the data and conclusions obtained were biased and factorial interactions were not clearly examined[7].This research work contains the Taguchi experimentation approach to design several rounds of experiments following the sequential experimentation strategy [10-11] for an in-depth discussion of the strategy. Therefore, the impact of each individual factor and factor interactions on surface roughness are clearly examined with a reasonably small amount of time and cost. Secondly, with the improved accuracy of today's machine tools and surface roughness measuring devices with the help of computers and software, the research work is able to include more parameters simultaneously with more accurate experimental data[12].

II. METHODOLOGY

Al-SiC MMC workpeice specimen having aluminum alloy 6061as the matrix and containing 15% vol of silicon carbide particles of mean diameter 25μm in the form of cylindrical bars of length 120mm and diameter Φ40mm was manufactured at vikram sarbhai space center Trivandrum by stir casting process with pouring temperature 700-710°C, Stirring rate 195rpm, extrusion at 457°C, extrusion ratio 30:1, direct extrusion speed 6.1m/min to produce Φ40mm cylindrical bars. The specimens were solution treated after 2 hours at a temperature of 540°C in a muffle furnace, Temperature were accurate to within ±2°C and quench delays in all cases were within 20 seconds. after solutionising, the samples were water quenched to room temperature, and subsequently aged for six different times to obtain samples with different Brinelll Hardness number(BHN), out of which two samples where selected, one with 94BHN obtained at peak age condition i.e. 2 hours at 220°C and other with overage condition i.e. 24 hours at 220°C respectively. All aged and solutionnized samples were kept in a refrigerator right after the heat treatments. In order to observe the effect of matrix hardness on turning of the composite materials with steam, compressed air, water vapor as coolant and dry cutting four samples has been selected. The selected material was manufactured by stir casting process[12-13]. As the matrix materials 99.9% pure aluminum was used, while 15 vol.% SiC particles

with an average size of 25μm were applied as the reinforcement element. The chemical composition of specimens given in Table.1

Element	Cu	Mg	Si	Cr	Al
Weight percentage	0.25	1.0	0.6	0.25	Balance

Nominal chemical composition of Base metal (6061 Al alloy): Table 1

The experimental study was carried out in panther lathe(2.5KW) for turning machining process. Cubic boron nitride(CBN) inserts KB-90(ISO code) are used as cutting tool for machining of MMC materials. The ISO codes of cutting tool insert and tool holder were shown in Table 2.

Tool holder specification	STGCR 2020K-16
Tool geometry specification	Approach angle:91° Tool nose radious:0.4 mm, Rake angle:0° Clearance angle:7°
Tool insert CBN(KB-90) specification	TPGN160304-LS

Details of cutting tool and tooling system used for experimentation: Table 2

The selected machining condition is given in Table 3. Surface condition of machined work piece was observed using JEOL JSM-6380LA analytical scanning electron microscope. Surface roughness was measured using Taylor/Hobson surtronic 3+ surface roughness measuring instruments.

Condition of machining	Turning
Machine tool used	Panther lathe(2.5KW)
Cutting speed(m/min)	150 m/min
Feed(mm/rev)	0.2 mm/rev
Depth of cut(mm)	0.5mm,1mm,1.5mm,2mm
Coolant used	Water vapor
Coolant pressure(Mpa)	0.7Mpa
Cooling distance(mm)	30mm

Machining conditions: Table 3

III.   EXPERIMENT SETUP

The water vapor generator and vapor feeding system are developed in which jet flow parameters (pressure, temperature, flow velocity and humidity) and cooling distance (it is the distance between nozzle and cutting zone) are controllable.

Figure.1 shows the principle of mechanism and Figure. 2 shows the water vapor generator and vapor feeding system[8].
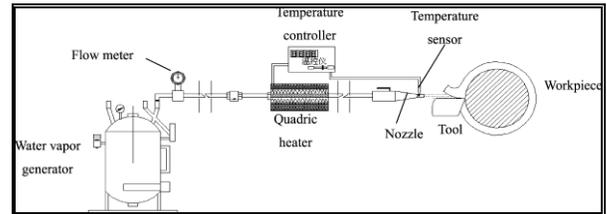
The principle skeleton of vapor gene:rator device and vapor feeding System: Figure 1
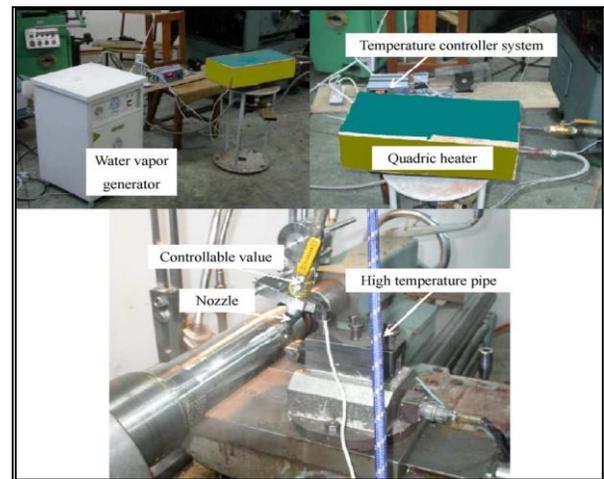
Vapor generator and vapor feeding system: Figure 2

A) Experimental details
Taguchi method was used for the execution of the plan of experiments, with three factors at three levels, The factors to be studied and the attribution of the respective levels are indicated in Table 4. The chosen array was L9 , which has 3 rows and 3 columns. The plan of the experiments consists of 9 tests[10-11].

B) Taguchi design of experiments
Taguchi orthogonal Array design.
L9(3*3)
Factors:3
Runs:9

Orthogonal array for Taguchi Design L9: Table 4

Speed (mm/rev)	Feed (mm/min)	Doc (mm)	Surface roughness(μm)
50	0.1	0.5	3.52
50	0.2	1.0	2.363
50	0.3	1.5	2.893
100	0.1	1.0	2.783
100	0.2	1.5	2.677
100	0.3	0.5	2.297

150	0.1	1.5	0.877
150	0.2	0.5	0.873
150	0.3	1.0	0.75

IV. PROBLEM OF OPTIMIZATION OF CUTTING CONDITIONS

The cutting parameters must be so selected that the machine is utilized to the maximum possible extent and that the tool life is as long as possible, when there are two conflicting objectives, a compromise must be reached. In general, the selection of easier operating conditions is not economically justified. If the cutting speed, feed and cutting depth are decreased, the work efficiency is reduced and the tool resistance to wear is prolonged[16]. In this way, the tool life is increased and the cost of the tool replacement reduced, but the labor costs are increased. Inversely, it is not always our aim to produce as much as possible within the shortest possible time. When selecting the optimum cutting conditions for some machine operation, we make a compromise between maximum material removal rate and the minimum tool wear[17-18]. The purpose of the optimization is to determine such a set of the cutting conditions v (cutting speed), f (federate), a (depth of cut) that satisfy the limiting equations and balances the conflicting objectives. The operation of turning is defined as a multiple-objective optimization problem with limiting non-equations and with three conflicting objectives (production rate, operation cost, and quality of machining). All the above-mentioned objectives are represented as a function of the cutting speed, feed rate and depth of cutting.

A) Objectives of Optimization
1. Production rate: Usually, the production rate is measured in terms of the time necessary for the manufacture of a product (T_p). It is the function of the metal removal rate (MRR) and of the tool life (T) [19]:

$$T_p = T_s + V \frac{\left(1 + T_c/T\right)}{MRR} + T_i, \qquad \text{...(1)}$$

Where T_s, T_c, T_i and V are the tool set-up time, the tool change time, the time during which the tool does not cut and the volume of the removed metal respectively. In some operations, the T_s, T_c, T_i and V are constants so that T_p is the function of MRR and T.

- The MRR: MRR can be expressed by analytical derivation as the product of the cutting speed, feed and depth of cut:

$$\text{MRR} = 1000vfa \qquad \text{.... (2)}$$

- Tool life (T): The tool life is measured as the average time between the tool changes or tool sharpening. The relation between the tool life and the related parameters is expressed by the well-known Taylor's formula:

$$T = \frac{k_T}{v^{a_1} f^{a_2} a^{a_3}}, \qquad \text{.... (3)}$$

where k_T, α_1, α_2 and α_3, which are always positive constant parameters and are determined statistically[20-21].

2. Operation cost: The operation cost can be expressed as the cost per product (C_p). In the cost of the operation, two values connected with the cutting parameters (T, T_p) [22] are distinguished:

$$C_p = T_p \left(\frac{C_1}{T} + C_l + C_o \right) \qquad \text{.....(4)}$$

where C_t, C_l and C_o are the tool cost, the labor cost and the overhead cost, respectively. In some operations, C_t, C_l and C_o are independent of the cutting parameters.

3. Quality of machining: The most important criterion for the assessment of the surface quality is roughness calculated according to

$$R_a = kv^{x_1} f^{x_2} a^{x_3} \qquad \text{.....(5)}$$

where x_1, x_2, x_3 and k are the constants relevant to a specific tool-work piece combination. In the presence of many incomparable and conflicting objectives, the ideal solutions satisfying all requirements are very rare. In order to ensure the evaluation of mutual influences and the effects between the objectives and to be able to obtain an overall survey of the manufacturer's value system, it is recommendable to determine the multi-attribute function of the manufacturer (y) [23-24] representing the company's/manufacturer's overall preference. A multiattribute value function is defined as a real –valued function that assigns a real value to each multi-attribute alternative, according to the decision maker's preferential order, such that more preferable alternative is associated with a larger value index than less preferable alternative. One global approach for the determination of the most desirable cutting parameters is by maximization of the manufacturer's implicit multi-attribute function.

B) Limitations
There are several factors limiting the cutting parameters. Those factors originate usually from technical specifications and organizational considerations. The following limitations are taken into account:
- Permissible range of cutting conditions: Due to the limitations on the machine and cutting tool and due to the safety of machining, the cutting parameters are limited with the bottom and top permissible limit.

$$v_{min} \le v \le v_{max}, \quad f_{min} \le f \le f_{max}, \quad a_{min} \le a \le_{max}$$

- Implied limitations arising from the tool characteristics and the machine capacity: For the selected tool, the tool maker specifies the limitations of the cutting conditions. The limitations on the machine are the cutting power and the cutting force. Similarly, the machining characteristics of the work piece material are determined by physical properties.
- Cutting power and force: The consumption of the power can be expressed as the function of the cutting force and cutting speed.

$$P = Fv/6122.45\,\eta \quad …. \qquad (6)$$

Where η the mechanical efficiency of the machine and F is is given by the following formula

$$F = k_n\, f^{\beta 1} d^{\beta 2} \quad \text{……..} \qquad (7)$$
$$P = k_n\, f^{\beta 1} d^{\beta 2}, \text{ where } k_n = k_f / 6122.45\eta$$

The problem of the optimization of cutting parameters can be formulated as the following multi-objective optimization problem:

min subject to limitations.

$$T_p(v,f,a), \min C_p(v,f,a), \min R_a(v,f,a)$$

The limitations of the power and cutting power and cutting force are equal to

$P(v,f,a) \le p_{max}, F(v, f, a) \le F_{max}$
   Force $\le$ 500 N
   Power $\le$ 2.5 Kw
Surface roughness $\le$ 2.5 µm

V. WORKING OF GA's

In GAs each variable is treated as a binary string corresponding to a gene. The variable set constitutes an individual, codified in a structure like the chromosomal one, having the genes one next to the other and more individuals constitute a population. In some cases decimal strings are used instead of binary ones, with the advantage of having strings with decimal ciphers much similar between them for two values near to one another. The population evolves owing to the modifications performed by the operators of crossover (interchange of chromosome segments between mating pairs) and mutation (variation of bits). Different strategies can be employed in the GAs and their efficiency can depend on the analyzed problem. On the basis of the efficiency of each individual, evaluated by a fitness, the genetic operator of selection chooses the good individuals, based on the principle of 'survival of the fittest', and are destined to the generation of a new population, by using both the genetic operators. Few worse individuals are destined to be modified deeply for the possible random change of all their genes[22-25].

Like it is known, the next generations have new characteristics, which can produce a better solution and however can favour the exploration of the feasible domain, reducing the risk of obtaining only local optima, with respect to traditional algorithms. Particularly the mutation on the worse individuals allows to renew the individuals destined to extinction, not dispersing their genetic patrimony, and, at the same time, increasing the diversity in the population and thus favouring the exploration of the design domain.

The employed strategy involves also the transfer of the best individual of each population into the next generation without transformations, replacing the worse one. Since for problems with few individuals, the best individual is usually transferred, it is believed that the higher the individual number, the higher must be the number of the transferred copies, replacing as many ones extracted randomly, in order to increase the possibility to enhance the population quality and to make the analyses faster; obviously the copy number must not be too high, in order to avoid that the solution tends to get stuck at a local optimum.

The process of going from the current population to the next population constitutes one generation in the evolution process of a genetic algorithm. Naturally, like in other optimization algorithms, the process is halted when the fitness stops to improve or a prefixed fitness has been achieved or the maximum iteration number has been reached.
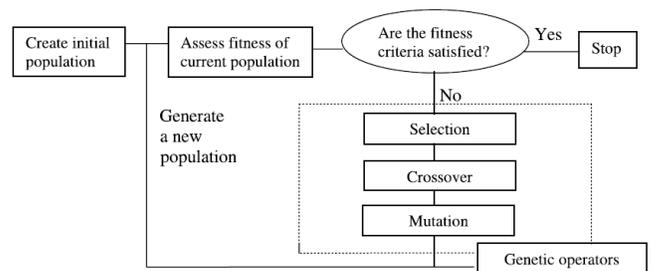
Flow chart of the basic genetic algorithm: Figure 3

A) GA results:

GA parameters: Table 5

Genetic Algorithm Values	
Population size	100
Total no of generation	200
Cross over Probability	0.70
Mutation probability	0.03

Total string length	24
Number of variables	3
Total runs to be performed	2

Lower and Upper bounds: Table 6

Speed(m/min)		Feed(mm/rev)		Depth of cut(mm)	
Min	Max	Min	Max	Min	Max
50	150	0.1	0.4	0.5	2.0

Out put Results cost of Production (Rs): Table 7

Best ever fitness	12.583372(From generation - 149)
Average fitness	13.77505
Worst ever fitness	70.92124

Optimized cutting parameters: Table 8

Speed(m/min)	Feed(mm/rev)	Depth of cut(mm)
70	0.102	1.617
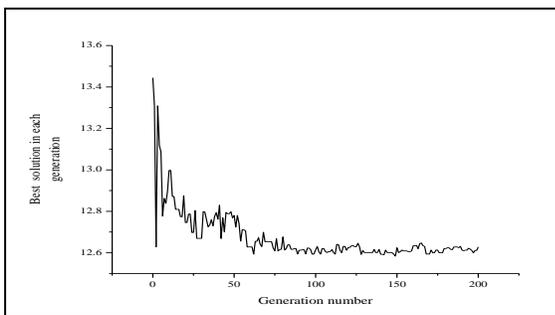
Number of feasible solution in each generation Vs Generation Number: Figure 4
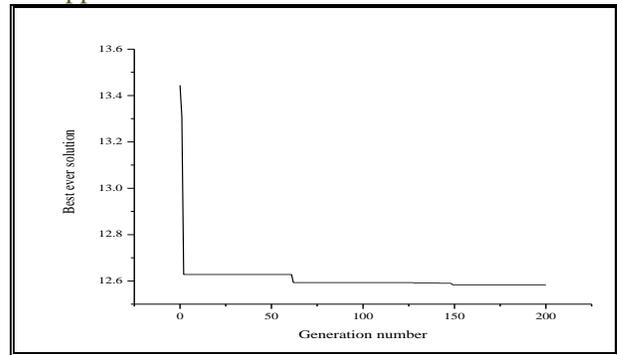
Generation Number Vs Best ever solution: Figure 5.

VI. CONCLUSION

Based on the GA parameter selected, it is observed from the figure 4, that best solution for cost optimization is observed at generation number 149 and is almost consistent there onwards (figure 5). These results are based on range of input machining parameters chosen. The same logic can be generalized to manufacturing environment by appropriate selection of range values for the input parameters. The involvement of surface response methodology and GA based optimization leads to an effective method of determining the process parameter values to achieve surface quality operational cost or production rate.

REFERENCES

[1] C. Ri-yao, *Metal Cutting Principle*, Machine Industry Publications China, China, 1984.

[2] G Boothroyd,. and W.A Knight, *Fundamentals of Machining and Machine Tool*. pp.141-200, Marcel Dekker, New York, 1989.

[3] Box, G. E. P., Hunter, W. G., and Hunter, J. S. *Statistics for Experimenters:Introduction to Design, Data Analysis, and Modeling Building*. John Wiley & Sons, New York, 1978.

[4] Box, G. E. P. and Draper, N. R., *Empirical Model-Building and Response Surfaces,* John Wiley & Sons, New York, 1987.

[5] T.G Dawson, T.R Kurfes, An investigation of tool wear and surface quality in hard turning, *Trans International jounal of Advanced Manufacturing* , SME 28,pp215–220,2000.

[6] A.Bhattacharya, Faria- R.Gonzalez, and I.Ham, Regression analysis for predicting surface finish and its application in the determination of optimum machining conditions". *ASME Journal of Engineering for Industry, 4,* pp. 711 – 714,1970.

[7] N.R Draper, and H.Smith, *Applied Regression Analysis*, pp.253-270, 3rd Edition. John Wiley & Sons, New York, 1998.

[8] Godlevski V.A., et al. ,*Water steam lubrication during machining Tribologia*, *vol. 162*, No. 6, 1998, 11, p. 890–901.

[9] V.A. Godleviski, A. V. Volkov, V. N. Latyshev, L. N. Maurine,*The kinetics of lubricant penetration action during machining*, *Lubrication Science* 9(2), pp.127-140, 1997.

[10] Montgomery, D. C.,*Design and Analysis of Experiments*, 5th Edition. John Wiley & Sons, New York. 2001.

[11] J.T Emanuel, and M.Palanisamy, *Sequential experimentation using two—level fractional factorials,* Quality Engineering, 13(3), pp. 335 – 346,2000.

[12] Feng, An experimental study of the effect of turning parameters on surface roughness in finish turning, *Proceedings of the 2001 Industrial Engineering Research Conference, Institute of Industrial Engineers, Norcross, GA*,2001.

[13] Jang, D. Y., Choi, Y. G., Kim, H.G., and Hsiao, A.,Study of the correlation between surface roughness measuring technique in hard turning, *International Journal of Machine Tools and Manufacturing*, 36(4) ,pp. 453-464,1996.

[14] *Minitab Meet MINITAB*, Release 13. Minitab Inc., State College, PA, 2000.

[15] D. W .Coit, B. T.Jackson, and A. E Smith,. Static neural network process models: considerations and case studies, *International Journal of Production Research, 36(11),* pp.2953 – 2967, 1998.

[16] S.K choudhury, I.V.K.Apparao, optimization of cutting parameters for maximizing tool life, *International journal of machine tool and manufacture 39,* pp.343-353, 1999.

[17] Feng, C-X. and Wang, X-F., Development of empirical models for surface roughness prediction in finish turning, *International Journal of Advanced Manufacturing Technology* . 2002.

[18] Rasch, F. O., and Rolstadas, A., Selection of optimum speeds and feeds in finish Turning, *Annals of CIRP*, 19, 787-792 , 1971.

[19] Phillips DT, Beightler CS.,Optimization in tool engineering using geometric programming. *AIIE Trans* 1970:2:3:355 – 60.

[20] Philipson RH, Ravindram A.,Application of mathematical programming to metal cutting. *Math program stud 1979*; 11:116 – 34.

[21] Nain CY, Yang WH, Tarng YS., Optimization of turning operations with multiple performance characteristic,*Journal of Material Process Technolgy,* 1999;95:90-6.

[22] Lee BY, Tarng YS.,Cutting – parameter selection for maximizing production rate or minimizing production cost inmultistage turning operations, *Journal of Material Process Technology*, 2000; 105 (7):61-6.

[23] Liu Y, Wang C., Neural network based adaptive control and optimization in the milling process, *International Journal of Advanced Manufacturing Technology*, 1990; 15:791 – 5.

[24] J.Kopac. M Bahor, M. Sokovic, Optimal machining parameters for achieving the desired surface roughness in fine turning of cold preformed steel workpieces, *International Journal of Machine Tools and Manufacture,* 42, 707 – 716,2002.

[25] G. Chryssolouris, Turning of hardened steels using CBN tools, *Journal of Applied Metal Working 2* ,100 – 106,1982.

[26] W. Grzeik, A revised model for predicting surface roughness in turning, *Journal of Applied Metal Working 2, 194*, 143 – 148, 1996.

[27] Sundaram, R. M. and Lambert, B. K., Surface roughness variability of ANSI 4140 steel in fine turning using carbide tools, *International Journal of Production Research, 17(3),* 249-258, 1979.

[28] Sundaram, R. M. and Lambert, B. K, Mathematical models to predict surface finish in fine turning of steel, Parts I and II, *International Journal of Production Research*, 19, 547-564, 1981.

[29] Abrao, A.M. and Aspinwall, D.K., The surface Integrity of Turned and Ground Hardened Bearing Steel; *Wear, Vol. 196*, pp. 279-284, 1996.

[30] Agha, S.R. and Liu, C.R., Experimental study on the performance of superfinish hard turned surfaces in rolling contact; *Wear 244*. pp. 52-59, 2000.

[31] Arsecularatne, J.A., Mathew, P. and Oxley, P.L.B.,Prediction of Chip Flow Direction and Cutting Forces in Oblique Machining with Nose Radius Tools"; *Proc. Inst. Mech. Engrs., Vol209(B)*, pp 305-315, 1995.

[32] Hasegawa, H., Seireg, A., and Lindberg, R. A., Surface roughness model for Turning, *International Journal of Advanced Manufacturing Technology* December, 285-289, 1976.

Challenges of Planning for Heritage Areas in Panaji City

Shaikh Ali Ahmed[1] and Dr. B. Shankar[2]

[1](Planning Assistant, North Goa Planning Authority) Goa
[2](Associate Professor in Urban and Regional Planning¸ Institute of Development Studies¸ University of Mysore) Mysore

ABSTRACT

Panaji is well known for its heritage and rich culture. Panaji was formally a Portuguese colony and now it is one of the important tourist destinations in India having one of the biggest coastlines. Panaji has been declared as Heritage city under redevelopment scheme covered by JNNURM by State and Central Government. The Outline Development Plan 2011 of Panaji has identified as many as 40 sites, monuments, buildings and houses which are of historical and architectural significance and listed in "The Goa Land Development and Building Construction Regulation 2010". There are many more heritage structures which needs identification and conservation. These structures have come under tremendous threat of new development and getting demolished due to non comprehensive approach by the agencies involved in conserving including the people at large. The paper focuses on the issues and challenges that are faced by heritage areas and planning measures for harmonious development of heritage areas in the City of Panaji.

Key Words: Heritage Areas, Conservation Areas, Challenges, Measures.

I. INTRODUCTION

Panaji is the capital city of the state of Goa, and headquarters of the North Goa district. Also, a historical city known for its rich culture, architecture and built heritage situated on the bank of Mandovi estuary. Panaji is the important centre in terms of Indo-Portuguese cultural heritage, having a number of natural, built, tangible and intangible sites and monuments. The City of Panaji has been declared as heritage city by the state and central Govt. It is spotted as one of the most attractive tourist destination and heritage centres in India, and also attracts a number of national and international tourist every.

Panaji is located in 15.25° North Latitude, 73.5° to East Latitude and about 60 Meter above the Mean Sea Level. Most of the important heritage churches are located in Old Goa, on the east of Panaji, about a distance of 10 kms. The unique cultural atmosphere is the result of the long absorbed 450 years of Portuguese rule. The influence has left a deep impact on the local traditions in all spheres and has formed a distinct cultural identity of the people. St. Francis Xavier was an instrument to carry with him the gospel of Jesus but more than that he also carried a way of life of people, their ethos and a rich culture in which

perhaps dance, music, arts, crafts, architecture, festivals, fun and frolic were all entwined with each other. Goa, therefore, has become "A symbol of religious co-existence, tolerance, and a unique example of cross fertilization of cultures in our vast, multi-racial, multi-lingual, multi-dimensional complex society". It developed a strong identity of its own for which the people are really proud of.

II. HISTORICAL BACKGROUND OF PANAJI

Panaji had humble beginnings as a part of Taleigao village. At that time it was little more than a large coconut grove bordered by paddy fields, with a few backwater ponds, canals and creeks. The original name of the town was Pahajani Khali, the land of backwaters. The official name is Panaji, though in the local language (Konkani) the Panaji gets pronounced as Ponnji, Ponnje, or Ponjhe. The Portuguese renamed Panajim as Panaji or Ponji. Panaji is a tiny city that packs in a large punch, built around a church facing a prominent square. The river Mandovi and the hillock of Altinho have historically been the determining factors for the city. During the 3rd century BC, Panaji and the rest of Goa were part of the Mauryan Empire. Panaji was the capital and principle city of Portuguese in 1843 after the collapse of Old Goa. It was thus a Eastern Empire, internationally renowned for its rich heritage, culture, famous monuments viz. churches, convents, temples, and beaches visited by hundreds of thousands of foreign and domestic tourists each year and one of the most popular holiday destinations, was also known as the Rome of the East. Panaji means "the land that never floods," was now renamed as Panjim by the Portuguese and it was also referred to as New Goa.

The city is bound by the Rua de Ourem creek on the north that has been paddy fields, with a few backwater ponds, canals and creeks. The original name of the town was Pahajani Khali, the land of backwaters artificially trained to flow along its eastern side, the Mandovi estuary on the north, the hillock of Altinho on the south east and the St. Inez nallah and Taleigao on the west.

III. HERITAGE BUILDINGS AND AREAS IN PANAJI

Panaji is the capital city of the state of Goa having beautiful residential, Institutional buildings and houses of rich architectural and historical significance. The Outline Development Plan 2011 of Panaji has identified as many as 40 sites, monuments, buildings and houses which are of historical and architectural significance and listed in "The

Goa Land Development and Building Construction Regulation 2010".

The areas of historical significance in the city have been designated as Conservation / Preservation areas by the Authority and notified in the O.D.P. 2011 and Goa Land Development and Building Construction Regulation 2010. These areas can be easily identified due to their unique character, existence of monuments and structures having rich architectural significance. The Outline Development Plan 2011 of Panaji has declared five areas as "Conservation Zone", and marked as "F" in the plan. They are: (1) Campal, (2) Mandovi river fronts (3) Fontainhas, (4) Portais, (5) Fonduvem-Ribandar. The total area comprises of 62.00 hectares and works out to be 15% of the settlement area approximately of Panaji Municipal limit excluding the unbuildable slope. These Conservation areas have diverse and unique characteristics.
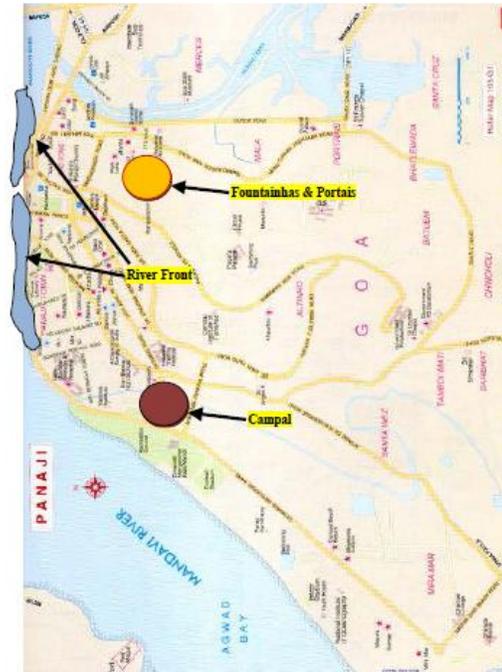
1. Fontainhas and Portais: Organic growth with Informal layout with streets and stepped accesses, high density mixed land use, interactive built-up and open spaces responding to human scale, although not of distinctive architectural style, the buildings together forms a cohesive group. In the absence of tree cover, buildings over looking narrow street/lanes provide shaded walkways and create an interesting play of lights and shadow.
2. Altinho: Predominantly institutional and residential use, individual large bungalows on undulating terrain framed by compound walls, ideally located facing the river front, devoid of through traffic and hence peaceful, low density of built up space.
3. Campal: Formal layout with Campal garden, facing the river front as focal point, individual low stung bungalows oriented towards riverfront, compound walls from strong visual element and provide distinctive character, green setting for built up areas adds to general ambience.
4. Mandovi River Front (Central Panaji): Orthogonal layout of streets creating well defined blocks with buildings facing the streets, internal open space within blocks correspond to larger city-spaces in the city, large number of state's historic public buildings with imposing facades that reinforce the city's identity. Low rise (G+1), Mangalore typed roof with repetitive elements. Predominantly commercial and institutional use with scattered residential buildings and extensive use of road space for social and cultural purposes (Ganesh festivals, Narkasur festivals, Church feasts, Carnivals etc.)

Fig. 1 Heritage Areas in Panaji

Fig.2 Modavi River Front

5. Fonduvem Ribandar: Predominantly residential and consists of old houses with fine architectural quality, having stepped and narrow accesses. The houses facing on either side of the road are fenced with solid compound walls. The rear strips of houses that are facing the river front are compounded with walls. There is slope for the area from kadamba hill to the old road leading to Old-Goa.

There is a great pressure of commercial development on the city's built heritage. High rise buildings have over shadowed heritage structures leading to aesthetic devaluation of the city. Natural heritage in terms of open spaces, trees lined avenues, historic precincts, and the river front are losing their functional and aesthetic identity.

IV. NEED FOR CONSERVATION

The city of Panaji is on the verge of getting developed from being part of fishing village of Taleigao in 1510 to a developed city in Goa. The evolution of the city has marked the achievement with the rise of many high rise buildings as a benchmark. The rapid growth in the economy especially during the 1990's, the city of Panaji could not resist from getting developed. It was during this period where all kinds of building activities emerged, colouring the city skyline, etc. Panaji is one of the cities that have its own history. The evolution of this city can be seen from the neoclassic architectural elements built during 1832. The culture and lifestyles are portrayed through the creation of built forms. This has made the old buildings significant in reflecting the culture and history that has formed the heritage of the city. The conservation of old buildings is a must in retaining the characters of the city. The density of new building designs seems not to be reflecting the true historical and unique character of the Panaji. The newly constructed buildings are modern in nature and don't reflect any aesthetics or architectural features. In contrast, the existence of old buildings, with their uniqueness, enhances the colourful urban form of the city of Panaji.

The city of Panaji is a fast growing city that the old heritage buildings are affected. The trend to develop the properties by the owners of the building has increased in which the economic factor play as the driven force. The prospective of getting big rewards by developing the land has resulted with the old buildings of significance and importance are getting demolished for new developments. The trends are irresistible since they could redevelop the city in par with the other developing cities without proper plans. The risk of losing the prospective incomes has put the owners not to retain the old buildings of high significance in its original forms. The purpose of the owners to develop the original heritage buildings, areas or houses in particular, is somehow in line with the motivated cost-effective factors as well as the desire to get rid of the old, decaying piece of architecture in absence of proper policy for maintenance of old heritage structures. Owners unknowingly the significance and importance of the heritage properties desires to develop the piece of land. It is the owner's right to develop the properties but could have been used as a mean to conserve the heritage properties.

Therefore the conservation areas need to be re-designated as "HERITAGE CONSERVATION" in the Outline Development Plan 2011 and for an effective integration or assimilation of the experience in the local context. The Government of Karnataka, Ahmadabad and Gujarat has been successful in bringing out the amendments for Heritage Conservation in their regulations. Educating the public at large and bringing the awareness of heritage monuments and areas should be put on top of the agenda. The high significance built heritage of Panaji needs to be conserved and calls for people's participation to protect the heritage areas from disorganised and unplanned development for retention of historical character and image of the city of Panaji.

V. ISSUES AND CHALLENGES

Panaji city has witnessed a tremendous change for the last two decades. Tall and disproportionate high rise buildings have over shadowed the heritage buildings. The heritage areas have almost lost their value and identity. This is due to the non availability of proper tools to control the heritage areas from disorganized development. The buildings of fine architectural features and aesthetic importance are demolished for the new development. Due to non availability of effective legislations and development control for the heritage areas, these areas needs to be re addressed in the context of conservation planning and development. The issues and challenges faced by the areas are as follows:

A) **Campal Area:**
- There is a great pressure for the new development in the Campal area due to its river front location.
- The new development lacks the aesthetics features, taller in height compared to the existing heritage structures, bulk in size and without any architectural features to suit the area.
- Private old heritage buildings that are of fine architectural features are being demolished giving way for the modern types of buildings due to lack of incentive programs and proper policy for maintenance and management of old houses of importance.
- Proper listing of the heritage resources is the need of the hour in the Panaji city.
- The Govt. owned buildings and Institutional buildings are maintained and managed.
- The new development seems to be not reflecting the true character of the heritage of Panaji city.
- The new development is not in harmony with the surrounding structures, aesthetics and in contravention to the height and number of floors.
- The new tall buildings will block the aesthetics vision of the heritage structures, there by losing the characteristics of the heritage area.

B) **The Mandovi River front (Central area):**
- The Central Mandovi River front area has developed rapidly due to its potential and being centrally located.
- High rise commercial buildings have changed the entire vision of the area.
- The buildings of aesthetic importance were sacrificed for the new development.
- High rise and high density buildings of stereo type have spoiled the beauty of Historical look of the city.

C) Altinho:

- The Development in the area so far has remained under control which is due to the area having undulating slope and extensively used by the elite and official residences of ministers and Govt. officials.
- The area towards the right side of Altinho slope has an organic form of settlement.

D) Fonduvem Ribandar:

- The very character of the area has remained unchanged, few new houses of double storey has been added to the area. Although the aesthetic is not in conformity with the surrounding houses the height and bulk of the new development is in par with the surrounding houses.
- Tall, high rise and high density buildings in the adjoining areas have over shadowed the heritage areas.

E) Fontainhas:

- The new development in the name of repairs and renovation cropping up in the area doesn't conform to the regulations in force and has spoiled the aesthetics of the area.
- Since there is no much scope for wider scale of development due to area being organic in nature, narrow width of road and on slope, it has remained undeveloped.

The Fontainhas area continues to be a focal point for Heritage tourists and travellers with a taste of past glories. Interest in Panaji's heritage buildings is worldwide. There are still clusters of quaint old houses in particular on one street in Portais and around St. Sebastian Chapel, which should be maintained. A sound policy towards the conservation is required and managed successfully.

Progress is inevitable and in generally a desired process but should be sustainable. The changes seen in the areas of heritage over the years are due to ineffective legislation and development control. The conservation Regulations 1989 has remained ineffective in controlling the heritage areas from disorganized developments. The present revised development control i.e., The Goa Land Development and Building Construction Regulation 2010 has also remained quiet in bringing out the strong and effective regulations for planned development of these areas. In view of the present revised regulations that allows more coverage and high FAR for smaller plots, the heritage areas will grow haphazardly like slums and without proper space in between the buildings and no areas will be left for parking purpose. Therefore, the need has aroused to readdress the issues and stop the areas from getting developed into an unplanned and disorganised manner.

VI. SWOT ANALYSIS

STRENGTHS	WEAKNESSES
1. Strategic location and well connection with other State capitals and other towns in the State. 2. An important Destination from heritage tourism point of view 3. A heritage town with a number of historical monuments, churches, Institutional buildings, houses and other places of interest. 4. Scope for expansion of major road networks 5. Availability of land for future expansion and development of the area. 6. Centre of attraction for nearby states, cities and villages.	1. Non availability of parking space in the city centre during the peak hours of the day. 2. Movement of vehicles and pedestrian in the main area is difficult due to traffic congestion. 3. Lack of frequent Public transport system in the city. 4. Limited civic, social and physical amenities 5. Disorganized development in the vicinity of the heritage area.
OPPORTUNITIES	THREATS
1. Connectivity is an important opportunity. 2. Riverfront development is in the city. 3. Tourism as a heritage tourism industry. 4. Enhancement of city's landscape and recreational spaces. 5. Area focuses on cultural and religious events at state and national levels. 6. Efforts to upgrade gardens and open spaces in the city and at neighbourhood level.	1. Uncontrolled building activities in the vicinity of heritage areas 2. Increase of migration forces due to availability of employment. 3. Changing socio-cultural ethos due to migration. 4. Construction on environmentally fragile area. 5. Chaotic development in the vicinity of heritage areas.

VII. PROPOSED MEASURES

The heritage resources of Panaji that are of high significant value should be properly protected, managed and transferred to the future generation in the context of sustainable development. Therefore these areas require conservative measures in protecting the rich heritage resources. These measures are identified and listed below:

i) There are many heritage resources which are not identified and it requires identification in each of the heritage areas for protection from getting demolished.

ii) Proper listing and inventory of the heritage resources are required to be done on war footings.

iii) The heritage areas need to be classified into different categories of conservation as per their historical development and uniqueness in characteristics.

iv) Diverse development control in each of the heritage areas is to be implemented for regulating these areas.

v) An incentive programme for protection and maintenance of the old heritage buildings/ houses should be carried out so that the burden of maintenance shouldn't be on the owner's shoulders.

vi) An heritage zones to be identified and a city specific heritage regulations are to be framed within the ambit of Town and Country Act of Goa State

vii) An awareness programme and importance of heritage structures and areas to be brought among the public at large.

viii) Insertion of transfer of development rights (TDR) in the development control for the heritage areas for those owners who lose their floor area (FAR) in conserving their properties in conservation zones.

ix) Effective and strong guidelines are required for regulating these areas from disorganised development.

VIII. CONCLUSION

The city of Panaji is a fast growing city that the old heritage buildings are affected. The trend to develop the properties by the owners of the building has increased in which the economic factor play as the driven force. It has rich in heritage resources, buildings, precincts and areas. The heritage areas are to be addressed in a comprehensive approach for harmonised development of heritage areas and reviving the traditional culture.

REFERENCES

[1]. Joshi C., et al, World Heritage Series Old Goa, Archaeological Survey of India, Goa.

[2]. Regional plan Goa, 2001 and 2021 prepared by Town & Country Planning Department, Government of Goa.

[3]. Report of the Sewri Consultants Pvt. Ltd., Mumbai, India and Sandhya Savant, Director Bombay Collaborative Urban design and Conservation Pvt. Ltd., Mumbai, India.on restoration of Capela da Nossa Senhora do Monte, Old Goa by Managing Director,

[4]. Report on JNNURM by STUP Consultants Pvt. Ltd. for Department of Tourism, Karnataka on Heritage and Urban renewal of Heritage core under J.N.NURM Scheme detailed

[5]. Government of Goa, The Ancient Monuments and Archeological Sites and Remains Act, No. 24 of 1958, Government of India.

[6]. Government of Goa, The Goa Ancient Monuments and Archeological Sites and Remains Act, 1978 and Rules, 1980.

[7]. Government of Goa, The Goa Town and Country Planning Act and Rules, Government of Goa, Planning and Development Authority (Development Plan), Regulations, 1989, Panjim Planning and Development Authority, Panjim.

1) Report of the Tourism Master Plan: GOA- 2011 final report prepared by Consulting Engg. Services (I) Ltd. for Government of Goa, Department of Tourism.

BIOGRAPHIES

Shaikh Ali Ahmed received M.Tech in Urban and Regional Planning from the University of Mysore, Mysore. Presently, he is working as Planning Assistant in North Goa Planning Authority, Panaji. He is Associate Member of the Institute of Town Planners,I ndia, He is presently pursing his Ph.D in Urban and Regional Planning at the Institute of Development Studies, University of Mysore. His research interests to include heritage conservation, heritage legislation.

Dr. B. Shankar received the B.E. degree in Civil Engineering in 1984, M.U.R.P degree in Urban and Regional Planning in 1989 and Ph.D degree in Urban and Regional Planning in 1997 from the University of Mysore, Mysore. He is working as Associate Professor in Urban and Regional Planning at the Institute of Development Studies, University of Mysore, Mysore. He was a Senior Faculty Member in the Administrative Training Institute and State Institute of Urban Development, Mysore, Government of Karnataka. He co-ordinated capacity building and training Programmes of Urban Poverty Alleviation Programmes of Karnataka State. His research interests to include Urban Planning, Urban Poverty, Community Development, Heritage Conservation, and Planning Legislation.

Perfect Domination Number and Perfect Bondage Number of Complete Grid Graph

Dr. D. K. Thakkar and D. D. Pandya
Department of Mathematics, Saurashtra university, Gujarat.

ABSTRACT
Grid graphs and domination are very important ideas in computer architecture and communication techniques. We present results about Perfect Domination Number and Perfect Bondage Number for Grid Graphs. We find Perfect dominating sets and Perfect domination number and Perfect bondage number for $G_{m,n}$ using special patterns.

Keywords: Grid graph, Complete grid graph, Domination number, Perfect Domination number, Perfect Bondage number.

1.INTRODUCTION
Starting in the eighties domination numbers of Cartesian products were intensively investigated. In the meantime, some papers on domination numbers of cardinal products of graph was initiated by Vizing[10]. He conjectured that the domination number of the Cartesian product of two graph is always greater than or equal to the product of the domination numbers of the two factors. This conjecture is still unproven. For complete grid graphs, i.e. graphs $P_k \times P_n$, algorithms were given for a fixed k which compute $\gamma(P_k \times P_n)$ in O(n) time [7]. In fact, the domination number problem for $k \times n$ grids, where k is fixed, has a constant time solution. In this paper we present a survey of Perfect domination numbers and Perfect bondage numbers of complete grid graphs $P_k \times P_n$, K=2,3,4,5.

2. DEFINITIONS
For notations and graph theory terminology, we follow Bondy and Murthy[1].
Let G=(V,E) be a simple graph with vertex set V and edge set E. A subset D of V is a dominating set of G if every vertex of V \ D is adjacent to at least one vertex of D. The minimum cardinality of a dominating set is called the domination number of G which is denoted as $\gamma(G)$.

A set D of vertices of a graph G is a Perfect dominating set if for every vertex u of V \ D ,

$|N(u) \cap D| = 1$. A Perfect dominating set of minimum cardinality is a minimum Perfect dominating set and its cardinality is Perfect domination number of G denoted by $\gamma_P(G)$.

The Perfect bondage number is minimum number of edges whose removal from original graph increase Perfect domination number in resultant graph denoted by $\beta_P(G)$.
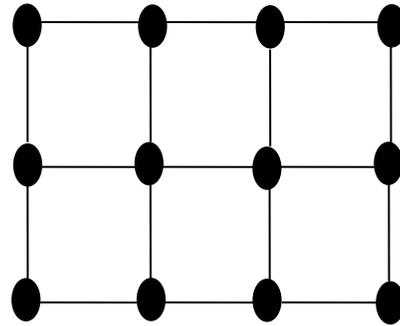
COMPLETE GRID GRAPH $G_{3,4}$

A two-dimensional complete grid graph is an $m \times n$ graph $G_{m,n} = P_m \times P_n$, the product of path graphs on m and n vertices. The Cartesian products of paths P_m and P_n with disjoint sets of vertices V_m and V_n and edge sets E_m and E_n is the graph with vertex set $V(P_m \times P_n)$ and edge set $E(P_m \times P_n)$ such that $((g_1,h_1),(g_2,h_2)) \in E(P_m \times P_n)$, if and only if either
$g_1 = g_2$ and $(h_1,h_2) \in E(P_n)$ or
$h_1 = h_2$ and $(g_1,g_2) \in E(P_m)$.
One example of a complete grid graph $G_{3,4} = P_3 \times P_4$ is shown in figure 1.

A grid graph $G_{m,n}$ has mn nodes and (m-1)n+(n-1)m=2mn-m-n edges. We observe that the path graph

$P_n = G_{1,n} = G_{n,1}$ and cycle graph $C_4 = G_{2,2}$.

From the definition of complete grid graph $P_k \times P_n$ we observe that for k=1 the grid graph is nothing but path graph that is $P_1 \times P_n = P_n \times P_1 = P_n$.

H-MERGE AND V-MERGE OPERATIONS:

From the definition of complete grid graph $P_k \times P_n$ we observe that for k=n=2 the grid graph $P_2 \times P_2$ is a cycle on 4 vertices.

H-Merging of two cycles

Let cycle C_1 and C_2 be two cycles on 4 vertices. Let cycle C_1 have vertex set $\{u_1, u_2, u_3, u_4\}$ and cycle C_2 have vertex set $\{v_1, v_2, v_3, v_4\}$ then vertex set of new graph obtained by H-merging denoted as $C_1{}^{\wedge}C_2$ is $\{u_1, u_2 = v_1, u_3, u_4 = v_3, v_2, v_4\}$. Edges in $C_1{}^{\wedge}C_2$ includes all the edges of C_1 and C_2 with $(u_2, u_4) = (v_1, v_3)$

This gives that

$|V(C_1{}^{\wedge}C_2)| = |V(C_1)| + |V(C_2)| - 2$ and
$|E(C_1{}^{\wedge}C_2)| = |E(C_1)| + |E(C_2)| - 1$

Thus H-Merging of two cycles gives a complete grid graph $G_{2,2}{}^{\wedge}G_{2,2} = P_2 \times P_3 = G_{2,3}$
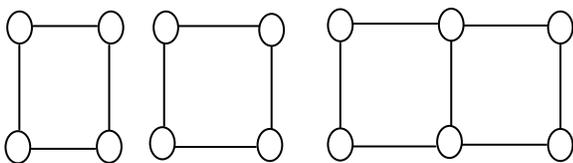
V-Merging of two cycles

In similar way we define another operation V-Merging which gives a complete grid graph shown as below.
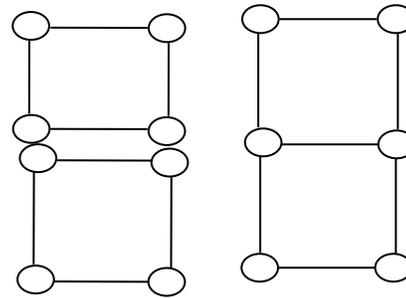
Lemma : 1

The Perfect domination number of grid graph $G_{2,n} = P_2 \times P_n$ for $n \geq 2$ is

$$\Upsilon_P(G_{2,n}) = \begin{cases} \left[\dfrac{n+1}{2}\right] & n = even \\[2ex] \left[\dfrac{n}{2}\right] & n = odd \end{cases}$$

Proof

Step : 1   Result is true for n = 2 $\Upsilon_P(G_{2,2}) = 2$
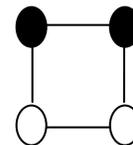
Step: 2 Suppose result is true for n = k, So,

$$\Upsilon_P(G_{2,k}) = \begin{cases} \left[\dfrac{k+1}{2}\right] & k = even \\[2ex] \left[\dfrac{k}{2}\right] & k = odd \end{cases}$$

Step : 3 If n = k + 1

CASE A: k=even so, $\Upsilon_P(P_2 \times P_k) = \left[\dfrac{k+1}{2}\right]$

n=k+1 is odd

To prove: $\Upsilon_P(P_2 \times P_{k+1}) = \left[\dfrac{k+1}{2}\right]$
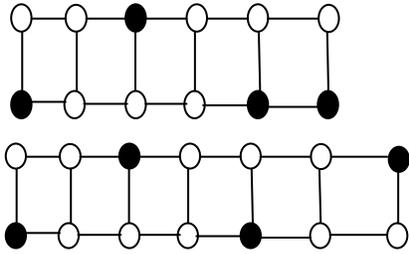
Consider $G_{2,k}$. Let D be Perfect dominating set in $G_{2,k}$. k=even So, In D we will have two adjacent vertices in last square given below.
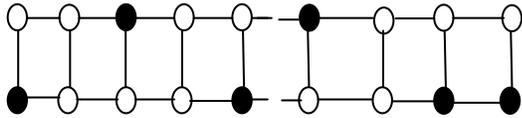
To construct $G_{2,k+1}$. We have to add 2 vertices(1 pair) in $G_{2,k}$. Now to perfect dominate $P_2 \times P_{k+1}$, we will do re-arrangement of D and with same number of vertices, we can perfect dominate $P_2 \times P_{k+1}$. See the figure given below.
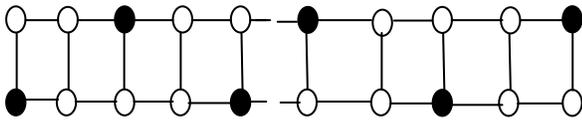
So, $\Upsilon_P(P_2 \times P_{k+1}) = \left\lceil \dfrac{k+1}{2} \right\rceil$

CASE:B k=odd

$$\Upsilon_P(P_2 \times P_k) = \left\lceil \dfrac{k}{2} \right\rceil$$

n=k+1 is even

To prove: $\Upsilon_P(P_2 \times P_{k+1}) = \left\lceil \dfrac{k+1}{2} \right\rceil$

Consier $P_2 \times P_k$, k=odd. Let D be perfect dominating set. We can observe that every vertex $\notin D$ is adjacent to only one

vertex $\in D$. see figure given below.
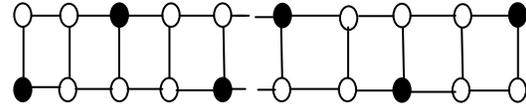
Now we add one pair in $P_2 \times P_k$ to construct $P_2 \times P_{k+1}$. In $P_2 \times P_k$, last pair is having one vertex $\in D$. So to perfect dominate $P_2 \times P_{k+1}$, we will have to add one more vertex.
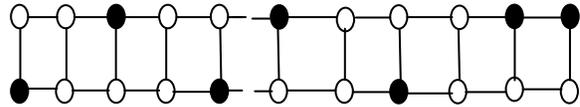
So, $\Upsilon_P(P_2 \times P_{k+1}) = \left\lceil \dfrac{k}{2} \right\rceil + 1 = \left\lceil \dfrac{k+1}{2} \right\rceil$

Lemma 2 The Perfect domination number of complete   grid   graph $G_{3,n} = P_3 \times P_n \ for \ n \geq 2$ is $\Upsilon_P(P_3 \times P_n) = n$

Proof:

Step : 1   Result is true for n = 2 $\Upsilon_P(P_3 \times P_2) = 2$
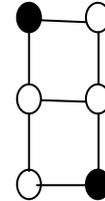
Step: 2 Suppose result is true for n = k, So,

$$\Upsilon_P(P_3 \times P_k) = k$$

Step : 3 If n = k + 1

To Prove: $\Upsilon_P(P_3 \times P_{k+1}) = k+1$

Consider $P_3 \times P_k$ and $\Upsilon_P(P_3 \times P_k) = k$

Let D be perfect dominating set with |D|=k

See following figure.
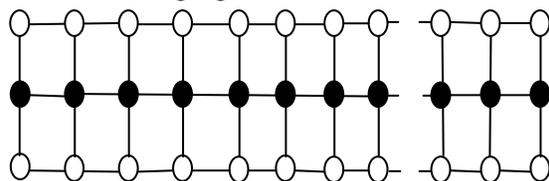
To construct P_{k+1}, we have to add $P_3 \times P_1$ at last pair. By figure given below

$$(P_3 \times P_{k+1}) = (P_3 \times P_k) \cup (P_3 \times P_1)$$
$$\Upsilon_P(P_3 \times P_{k+1}) = \Upsilon_P(P_3 \times P_k) + \Upsilon_P(P_3 \times P_1)$$
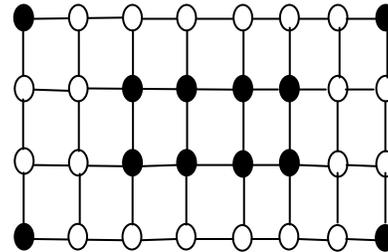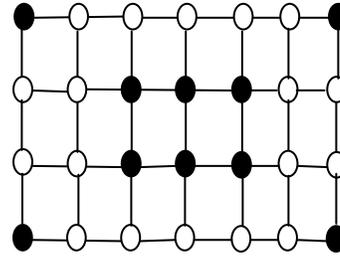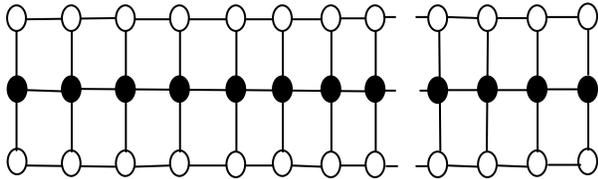
Lemma 3:- The Perfect domination number of grid graph $P_5 \times P_n$ for $n \geq 4$ is 2n-4

Proof :

Step:-1 Result is true for n=4

$$\Upsilon_2(P_4 \times P_4) = 4$$
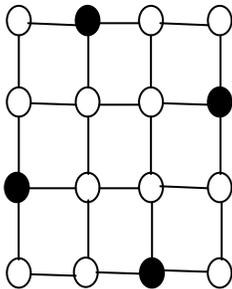
Step:-2 Suppose result is true for n=k

$$\Upsilon_2(P_4 \times P_k) = 2k - 4$$

Step:-3 Check result for n=k+1

To prove: $\Upsilon_2(P_4 \times P_{k+1}) = 2(k+1) - 4$

Consider one example of $P_4 \times P_7$ & $P_4 \times P_8$

Now consider in general form $P_4 \times P_k$
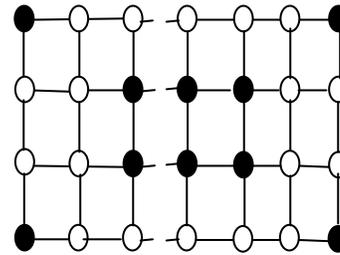
Let D be Perfect dominating set in $P_4 \times P_k$ with Cardinality 2K-4. Now consider $P_4 \times P_{k+1}$

To Construct $P_4 \times P_{k+1}$ we add one path $P_4 \times P_1$ at the end.
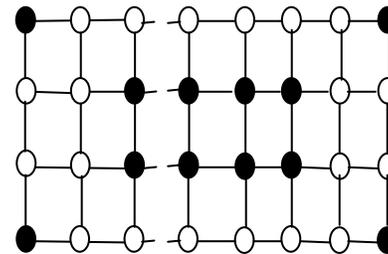
To dominate $P_4 \times P_{k+1}$, we have to add two more vertices in D (Observing pattern)

$$\Upsilon_P(P_4 \times P_{k+1}) = (2k-4)+2 = 2k-2 = 2(k+1)-4$$

Lemma 4: The Perfect domination number of grid graphs $P_5 \times P_n$ for $k \geq 3$ is

$$\frac{5n}{3} \quad ; \quad n = 3k$$

$$5\left(\frac{n+2}{3}\right) \quad ; n = 3k+1$$

$$5\left(\frac{n+1}{3}\right) \quad ; \quad n = 3k+2$$

Proof:

Case 1: if n=3k

Now divide the $5 \times n$ grid graphs into k blocks $P_5 \times P_3$ each. Extending the above result for $P_5 \times P_{3k}$ where each of the k blocks $P_5 \times P_3$ contribute 5 vertices into the minimal Perfect dominating set. We get 5k vertices into minimal Perfect dominating Set. Using the minimality of Perfect dominating set of $G_{5,3}$, we claim the minimality of Perfect dominating set of $P_5 \times P_{3k}$.

Hence $\Upsilon_P(P_5 \times P_{3k}) = 5k$

Case 2: if n=3k+1

We divide the $5 \times n$ grid graphs into k blocks of $P_5 \times P_3$ each and a path $P_5 \times P_1$. As in case 1 we get patterns in the k blocks of $P_5 \times P_3$ giving 5 vertices into the Perfect dominating set from each block for Perfect domination of the vertices in the block $P_5 \times P_1$, We need 5 vertices to be added to Perfect dominating set. Thus the Perfect dominating set of $P_5 \times P_{3k+1}$ has 5k+5 vertices. As in case 1 we claim that this is the minimal Perfect dominating set of $P_5 \times P_{3k+1}$.

Hence
$$\Upsilon_P(P_5 \times P_{3k+1}) = \Upsilon_P(P_5 \times P_{3k}) + \Upsilon_P(P_5) = 5k+5$$
$$= 5(k+1)$$

Case 3: if n=3k+2

We divide the $5 \times n$ grid graphs into k blocks of $P_5 \times P_3$ each and a path $P_5 \times P_2$. As in case 1 we get patterns in the k blocks of $P_5 \times P_3$ giving 5 vertices into the Perfect dominating set from each block for Perfect domination of the vertices in the block $P_5 \times P_2$, We need 5 vertices to be added to Perfect dominating set. Thus the Perfect dominating set of $P_5 \times P_{3k+2}$ has 5k+5 vertices. As in case 1 we claim that this is the minimal Perfect dominating set of $P_5 \times P_{3k+2}$.

Hence
$$\Upsilon_P(P_5 \times P_{3k+2}) = \Upsilon_P(P_5 \times P_{3k}) + \Upsilon_P(P_5 \times P_2) = 5k+5$$
$$= 5(k+1)$$
(Refer following figures.)
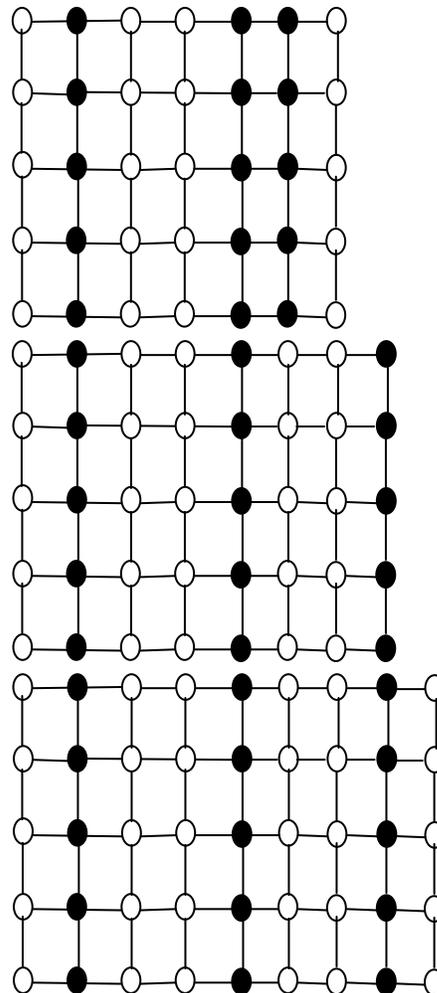
Lemma 5: The Perfect bondage number of complete grid graph $G_{2,n} = P_2 \times P_n$ for $n \geq 5$ is

$$\beta_P(P_2 \times P_n) = \begin{array}{ll} 1, & n = odd \\ 2, & n = even \end{array}$$

**Proof:**

Consider the complete grid graph $P_2 \times P_n$ .By lemma 1perfect domination number ,

$$G_{2,n} = P_2 \times P_n \ \ for \ n \geq 2 \ is$$

$$\Upsilon_P(G_{2,n}) = \begin{array}{ll} \left[\dfrac{n+1}{2} \right] & n = even \\ \\ \left[\dfrac{n}{2} \right] & n = odd \end{array}$$

Case 1. Let n=even

Let D be perfect dominating set in this graph.. Then by Lemma 1,
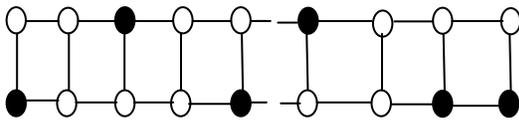
Now if we remove an edge between last pair then in resultant graph , Let S is perfect dominating Set. (see following pattern)

$| S |=| D | +1$
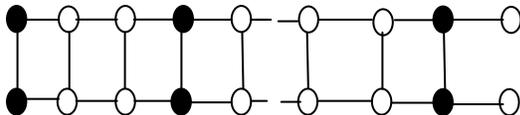
So, Perfect Bondage number for this graph is 1.

Case 2 Let n=odd

Let D be perfect dominating set in this graph.. Then by Lemma 1,
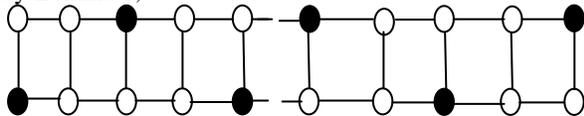
Now if we remove two edges from graph such that last square becomes isolate from graph and let S be perfect dominating set in resultant graph then
$| S |=| D | +1$
(see following pattern)
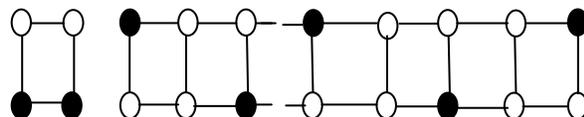
So, Perfect Bondage number for this graph is 2.

Lemma 6

The Perfect bondage number of complete grid graph

$$G_{3,n} = P_3 \times P_n = \begin{array}{ll} 2 \ , \ otherwise \\ 1 \ \ , \ \ \ n = 5,1 \end{array}$$

Proof:

By Lemma 2,
The Perfect domination number of complete grid graph $G_{3,n} = P_3 \times P_n$ for $n \geq 2$ is
$$\Upsilon_P(P_3 \times P_n) = n$$

Let D be Perfect Dominating Set in graph.Observe following figure.
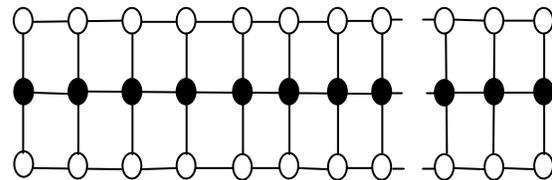
We can see that all corner vertices are not in D. If we make any corner vertex isolate by removing its two adjacent edges then Let S be perfect dominating set in resultant graph. See following figure.
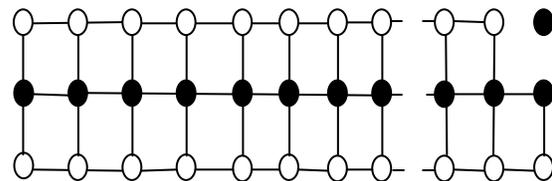
We can See that
$| S |=| D | +1$
So, Perfect Bondage number for this graph is 2.

Lemma: 7   The Perfect bondage number of complete grid graph $G_{4,n} = P_4 \times P_n$ is 2 $n \geq 2$

Proof:
Consider the complete grid graph $G_{4,n} = P_4 \times P_n$

By lemma 3, $\Upsilon_P(P_4 \times P_n) = 2n - 4$.

Now consider in general form $P_4 \times P_k$
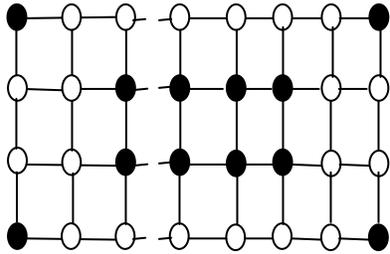
Let D be perfect dominating set in this graph.

Now if we remove two edges which are adjacent to corner vertex to make corner vertex isolated,then Let S be perfect dominating set in resultant graph.
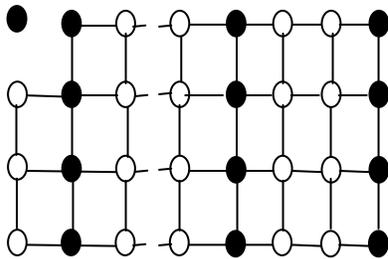
Then |D| > |S|

If n=3k+1 then |S|=|D|+3

Otherwise |S|=|D|+1

Here there exist not any other perfect dominating set having cardinality minimum then S.

So, Perfect bondage number of this graph is 2.

Lemma :8

The perfect bondage number of complete grid graph

$$G_{5,n} = P_5 \times P_n = 1, \qquad n \geq 5$$

Proof:

Consider the complete grid graph $P_5 \times P_n$.

For example $P_5 \times P_4$ and $P_5 \times P_5$
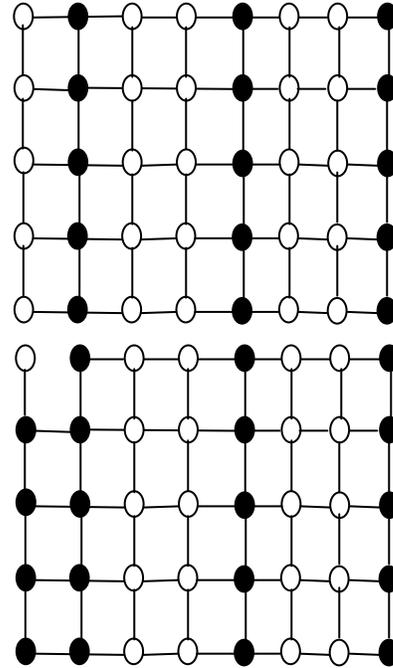
By lemma 4,

The Perfect domination number of grid graphs $P_5 \times P_n$ for $k \geq 3$ is

$$\frac{5n}{3} \qquad ; \qquad n = 3k$$

$$5\left(\frac{n+2}{3}\right) \quad ; n = 3k+1$$

$$5\left(\frac{n+1}{3}\right) \; ; \; n = 3k+2$$

Let D be perfect Dominating set in $P_5 \times P_n$.
Consider corner vertex if we remove an edge which is adjacent to any corner vertex $\notin D$ and having another end vertex $\in D$.
Then perfect Dominating number of resultant graph will increase.

Graph	Domination Number $\Upsilon(P_k \times P_n)$	Perfect Domination Number $\Upsilon_P(P_k \times P_n)$	Perfect Bondage Number $\beta_P(P_k \times P_n)$
$P_2 \times P_n$	$\left\lfloor \dfrac{n+2}{2} \right\rfloor, n \geq 1$	$\left\lceil \dfrac{n+1}{2} \right\rceil \quad n = even$ $\left\lceil \dfrac{n}{2} \right\rceil \quad n = odd$	1 if n=odd 2 if n=even
$P_3 \times P_n$	$\left\lfloor \dfrac{3n+4}{4} \right\rfloor, n \geq$	n	1 if n=5,1 2 otherwise
$P_4 \times P_n$	n+1 n=1,2,3,5,6,9 n otherwise for $n \geq 1$	$n \geq 4$ is 2n-4	2
$P_5 \times P_n$	$\left\lfloor \dfrac{6n+6}{5} \right\rfloor$ $n = 2, 3, 7$ $\left\lfloor \dfrac{6n+8}{5} \right\rfloor$ otherwise $n \geq 1$	$k \geq 3$ $\dfrac{5n}{3} \quad ; \quad n = 3k$ $5\left(\dfrac{n+2}{3}\right) \; ; n = 3k+1$ $5\left(\dfrac{n+1}{3}\right) \; ; \; n = 3k+2$	1

REFERENCES

1. Bondy and Murthy: **Graph Theory with applications**, Mac Millan, (1976)

2. Chang, T.Y.Clark, W.E.: **The domination numbers of the and the grid graphs. The journal of Graph Theory**, Volume 17, Issue 1, (1993), 81-107

3. Bondy and Murthy: **Graph Theory with applications**, Mac Millan, (1976)

4. Chang, T.Y.Clark, W.E.: **The domination numbers of the and the grid graphs. The journal of Graph Theory**, Volume 17, Issue 1, (1993), 81-107

5. Deo Narsingh: **Graph Theory with Applications to Engineering and Computer Science,** Prentice Hall, (2001)

6. Gary Chartrand Linda Eroh Frank Harary Ping Zhang: **How lagre can the domination numbers of a Graph be?** Australasian Journal of Combinatorics 21,(2000),23-35

7. Hare, E.O. **Algorithms for grid and grid-like graphs**, Ph. D. Thesis, Dept. Computer Scie. Clemson University, 1989.

8. Haynes, T.W.Hedetniemi, S.T.Slater, P.J.**:Fundamentals of Domination in graphs,** Marcel Dekker Inc. New york, (1998).

9. Bondy and Murthy: **Graph Theory with applications**, Mac Millan, (1976)

10. Chang, T.Y.Clark, W.E.: **The domination numbers of the and the grid graphs. The journal of Graph Theory**, Volume 17, Issue 1, (1993), 81-107

11. Deo Narsingh: **Graph Theory with Applications to Engineering and Computer Science,** Prentice Hall, (2001)

12. Gary Chartrand Linda Eroh Frank Harary Ping Zhang: **How lagre can the domination numbers of a Graph be?** Australasian Journal of Combinatorics 21,(2000),23-35

13. Hare, E.O. **Algorithms for grid and grid-like graphs**, Ph. D. Thesis, Dept. Computer Scie. Clemson University, 1989.

14. Haynes, T.W.Hedetniemi, S.T.Slater, P.J.**:Fundamentals of Domination in graphs,** Marcel Dekker Inc. New york, (1998).

15. Jacobson, M.S.Kinch, L.F.: **On the domination number of products of Graphs,** I,Ars Combinatoria, 18 (1983), 33-44

16. Jochen Harant Michael A.Henning: **On Double Domination in Graphs,** Discuss.Math.Graph Theory 25,(2005),29-34.

17. Tony yu chang, W. Edwin Clark, Eleanor O.Hare: **Domination Numbers of Complete Grid Graphs**, I. Ars Combinatoria, 69,(2003),97-108.

18. V.G.Vizing,: **The Cartesian Product of Graphs**, Vychisl. Sistemy 9 (1963), 30-43.

Study of Wind Catchers with square plan: Influence of physical parameters

Mahmud Dehnavi[1*], Maryam Hossein Ghadiri[2], Hossein Mohammadi[3], Mahdiar Hossein Ghadiri[4]

[1,2](Department of Architecture, Dezful Branch, Islamic Azad University,Dezful, Iran)
[3,4] (Department of Electronics, Dezful Branch, Islamic Azad University,Dezful, Iran)

ABSTRACT

This paper is a synopsis of the results of a research on the wind catcher, the cooling systems in traditional Iranian architecture which used in cities with hot-dry and hot-humid climates. This review demonstrates wind towers' characteristics with emphasis on their morphology. Different ratios between different elements of wind catcher such as its plan's length to width, its tower height to shelf height can be fundamental information in the design of wind catcher as a sustainable element on building. This paper is evaluate the performance of square wind catchers in order to find the most efficient form of square wind catcher in decreasing the indoor air temperature which is square wind catcher with plus blade form. Experimental results are evaluated by numerical method conducted by Fluent.

***Keywords* - Wind Catcher, Natural Cooling System, Wind catcher's Blade, Square plan, Fluent**

I. INTRODUCTION

Environmental and natural phenomena play a critical role in laying the region's interrelated cultural, economic and social infrastructures. Traditional buildings in the Iranian desert regions are constructed according to the specific climatic conditions and differ with those built in other climates. Due to lack of access to modern heating and cooling equipment in ancient times the architects were obliged to rely on natural energies to render the inside condition of the buildings pleasant. In the past, without modern facilities, it was only the intelligent architecture of the buildings that enabled people to tolerate the hot summer. A cross sectional view of wind tower is shown in figure 1.. It is seen in settlements in hot, hot-dry and hot-humid climates. They look like big chimneys in the sky line of ancient cities of Iran. They are vertical shafts with vents on top to lead desired wind to the interior spaces and provide thermal comfort. This architectural element shows the compatibility of architectural design with natural environment. It conserves energy and functions on the basis of sustainability principles [1].
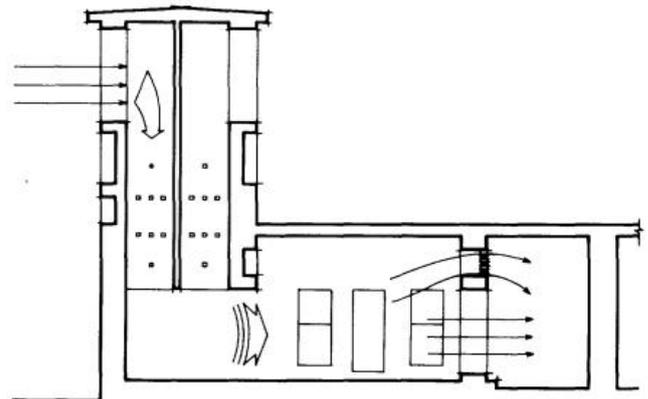

Fig. 1   Cross section of wind catcher

II. ORIENTATION AND FUNCTION OF WIND CATCHER

The orientation of wind tower generally means the positions of the wind tower flank based on the four main geographical directions. It is determined in view of function, use of wind power and the desired direction in which the wind blows. One-directional wind towers can be found in Meibod and they are positioned to face the desired wind. In some cases one directional wind tower acts as air suctioning mechanism with turned its back away from wind direction creating negative pressure which allows warm air of interior to flow out from the house. Desirable wind currents in Yazd blow from the north-west, therefore the long sides of wind towers are oriented north-west for maximum usage of the wind to provide cooling for buildings. In coastal regions like Bandar Lengeh, wind towers have an east-west orientation due to sea breezes and desirable winds that flow during daytime and nighttimes from that direction. Wind towers are also built with a four-directional orientation in order to use all of the desirable winds from north to south and from east to west. Orientations of the wind towers are based on the direction of the main desired wind.

A Wind tower is a formal structural element in Iranian architecture that is used to convey the wind current to the interior spaces of buildings in order to provide thermal comfort for occupants. In Iranian architecture a wind tower is a combination of inlet and outlet openings. A tunnel connected to wind tower provides cool air for building while also serving as a conduit through which the stuffiness within building is conveyed out through wind tower's shaft. There were wind towers in Bam which were destroyed by earthquakes and weren't directly connected to

the living hall. They were built away from the house and an additional underground tunnel links the base of the wind tower to the basement.

In most wind towers, especially the four sided types, the tower is divided by partitions. One of the shafts operates all the time to receive the breeze and the other three shafts work as outlet air passages. They convey the stuffiness out of the living space through the "flue" (chimney) effect. The chimney effect is based on the principle that the air density increases with the increase in temperature. The difference in temperature between the interior and exterior parts of a building creates different pressures and result in air currents. The average relative humidity or moisture in hot and dry regions is low and it is necessary to introduce humidity within wind tower's system through evaporative cooling. The air current which enters the wind tower is first passed over a stone pond and fountain, thereby bringing humidity to the interior spaces of building.

In some places, mats or thorns are placed within the wind tower and poured with water to increase humidity which helps cool the air flow. The hot weather in Yazd has the potential effect of causing water to evaporate easily to develop cooling in the living spaces and relative humidity in the air, thereby reducing the heat and dryness. It is clear that there is usually high humidity in hot and humid regions because of their vicinity to the sea. In these regions, wind towers reduce the temperature of the surrounding only through the movement of the air they facilitate, not through increased humidity. The level of humidity in this region is already high and an increase of humidity would create uncomfortable living condition. A wind tower in a hot and dry region brings about comfort by evaporation and air motion but a wind catcher in a hot humid region only moves the air and conveys the wind into spaces. Different function and shapes of wind towers were designed for different climates [2].

III. CRITERIA IN WIND CATCHERS PERFORMANCE AND DESIGN

A wind-catcher is an effective cooling device and constant structure in Iran architecture. It leads desirable wind through the inner part of the building to facilitate thermal comfort. There are actually two main principles in wind-catchers functioning:

1. The principle of traction for opening facing the wind and suction for openings facing away from the wind.

Wind-catcher serve dual functions; bringing fresh air inside building and sending hot and polluted air outside through' a suction process." When the wind hits against the walls of internal blades of the wind-catcher, airflow is brought inside the building. Inversely, when the wind-catcher's hole is turned away from the wind direction, the hot and polluted air will be sucked and released into the wind outside thus functioning as a ventilation or suction machine.

When the wind hits an obstacle, the density of the air is thick on the side of the wind direction, so in this direction there is a positive pressure, at the same time a negative pressure is created on the other side. Based on this principal, wind-catcher's opening which faces the wind,

takes in the air while its opening facing away from the wind direction will draw out air (Fig. 2).

2. The temperature difference

It seems little attention has been given to the technicalities and roles of temperature difference in wind-catchers. Wind-catcher's functioning relies on temperature difference when there is no sensible wind available.
During the day, since the sun hits on the southern face of the wind catcher, the air heats in the southern face of the wind catcher, and goes up. This air taken above through the inner air of the porch is compensated and in fact it makes a kind of proportional vacuum inside the porch, and takes the cool air of the inner court into itself, so the existing air in the northern opening is pulled down too (Fig. 3).

During the night it becomes cold outside, and the cold air moves down. This air is saved by the heat and becomes warm on parapets and then goes up. This circle continues till the temperature of the walls and outside temperature become equal. But before it usually arrives at this situation the night ends and once again the wind-catcher acts its function as mentioned above [3].
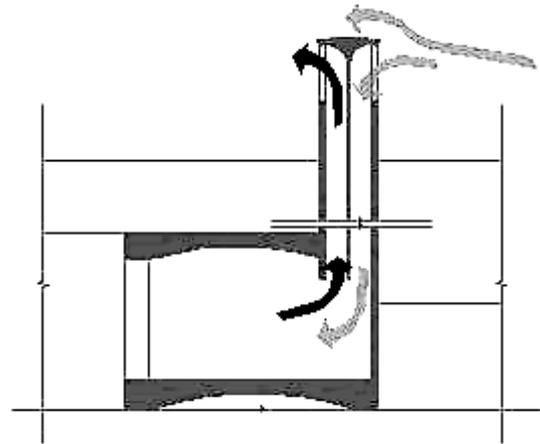

Fig. 2   traction and suction in wind-catcher.
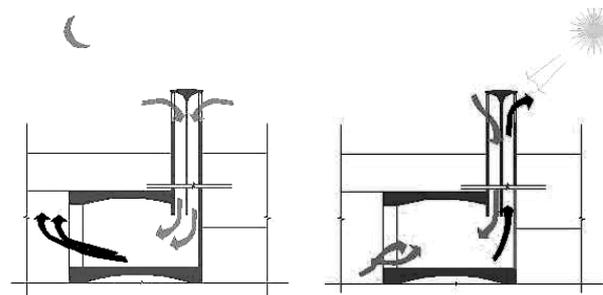
Fig. 3   Wind-catcher function during the day and night

IV. CATEGORIES OR TYPES OF WIND CATCHERS BASED ON ORIENTATIONS

Wind catchers in Yazd are categorized into five groups by roof, based on their direction.

1. The one directional towers

One directional towers (figure 4) generally face north-west or north. They have a sloping roof and one or two vents only. Otherwise they are commonly described by the

direction in which they face such as "Shomali" or north facing. Based on a survey by Roaf [4] 3% of the wind towers in Yazd were unidirectional [5].


Fig. 4   One directional towers plan

2.   The two directional towers

The tower is divided into two shafts by a vertical brick partition as shown in figure 5. It has only two vents. They are often called by direction, such as north-south towers. Roaf's survey indicates that 17% of the towers are of this kind in Yazd and all are found on the ordinary houses.

Fig. 5   Two directional towers plan

3.   Three-sided wind-catcher

This type of wind-catcher is not usual. A little number of this form can be found in Tabas.

4.   The four directional towers

Studies indicate that this is the most popular wind tower. They have four main vertical shafts divided by partitions. More than half of the wind towers in hot and dry region are of this type and they are called Yazdi. Almost all wind towers in hot humid region are the four sided type which is shown in figure 6[6].
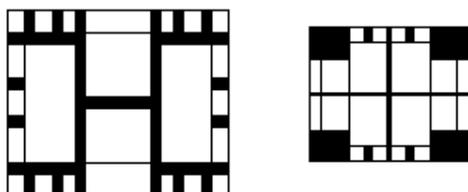

Fig. 6   Four directional towers plan

V. CATEGORIES OR TYPES BASED ON PLAN FORMS

Generally speaking, in Iran wind catchers have been recognized in a variety of forms and plans (figure 7) such as circle, Octagon, polygon, square and oblong. However, no triangular form has yet been recognized or located anywhere in the Middle East. Wind catcher with a circular plan or form is very rare and such type of wind catcher does not exist in Yazd. The square form is the type used in the four directional wind towers in Yazd. The rectangular forms consist of one, two, four directional wind towers.

Eight directional wind towers are those with an octagonal plan.

Fig.7   categorization of wind catchers

Partitions are component in wind towers which divide it into several shafts. They are built of mud brick. These partitions form a plane grid of vents ending to a heavy masonry roof on top of the tower. Partitions can be classified into main partition and secondary partitions [7]. Main partitions continue to the centre of the tower, forming a separate shaft behind the vents. These partitions often start between 1.5-2.5 m above the ground floor level. The patterns of the partitions vary from tower to tower, but the most commons are in forms of I, H and diagonal. Secondary partitions remain as wide as the external wall, about 20-25 cm. A shaft can be subdivided by a number additional partitions performing either structural or thermal role. These can separate the tower, respectively into two or four shafts. Partitions divide wind tower into small shafts to increase air motion according to "Bernoly effect". Air flow rate increases when air passes through narrow section [8]. Such an arrangement also provides more surfaces in contact with the flowing air, so that the air can interact thermally with the heat stored in the mass of these partitions. They function as fins of radiator because mud brick partitions give back stored heat during the night and are also good in absorbing heat. Contact between warm wind and mud brick partitions can transfer heat to the partitions, thus wind with less heat will enter the space.

VI. EXPERIMENT

Four square form wind catchers with different blades form including +, x, H and K which are shown in figure 8 as model 1 to 4 were investigated to find out the effect of wind catchers on decreasing the indoor air temperature. The air temperature inside and outside the wind catchers are measured by a thermometer. The dates selected for the experiments is on the 23rd and 24th of July which represent the hottest days in Yazd. In order to comparison all selected wind catchers have almost same specification such as tower

height, shelf height, plan dimension and wind catcher room size which is represented in table 1. The only variable in this experiment is the blade form of wind catchers.
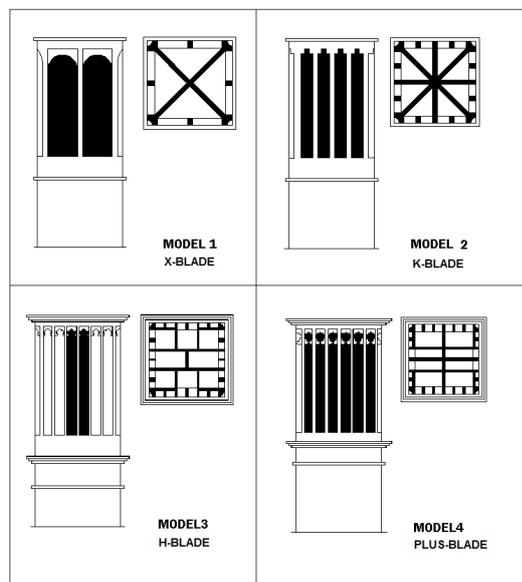

Fig. 8   Plan and elevation of investigated models

Table1. Specification of selected square wind catchers

Type of square wind catcher	Tower height (m)	Wind catcher room (m²)	Plan dimension (m²)	Shelf height
X blade	8.50	6×4	1.5× 1.5	1.50
K blade	8.35	6× 4	1.45× 1.45	1.45
H blade	8.42	5.8× 3.9	1.42× 1.42	1.42
+ blade	8.56	6×4	1.5×1.50	1.48

VII. RESULT

Through the chart in figure 9 when the outdoor temperature is maximize around 43ºC on 2:30 pm, the wind catcher room temperature is around 39 ºC in model 1, 38c in model 2 ,38.5 ºC in model 3 and 36 ºC in model 4.hence these kinds of wind catchers can decrease the temperature between 3ºC to 7ºC while outdoor temperature is in the highest level. Outdoor temperature decrease gradually in the afternoon while indoor air temperature decrease 3ºC to 5ºC up to 8pm. After 8 at night, indoor air temperature is 3ºC to 4ºC more than outdoor temperature, which is the time that courtyard air temperature is less than indoor air temperature and indoor thermal comfort could be obtain by opening the windows of room. As we can see through the chart, indoor air temperature in wind catcher room with plus form blade wind catcher is than other samples. As 36 ºC is not an ideal comfort temperature, ancient architects use water pond under the wind catcher for evaporative cooling in order to decrease the indoor air temperature. In order to validate the results numerical method is employed conducting by Fluent 6.3 for X-blade wind catcher which is shown in figure 10. As shown in figure 11, the difference between numerical result and experimental result is 8% which is acceptable.

VIII. CONCLUSION

The wind catcher is an intelligent exploitation of wind energy which makes possible thermal comfort in hot region. The major advantage of the wind towers is that they are passive systems, requiring no energy for their operation. Wind catchers can be categorized based on the ratios between their different parameters such as height, width, and length as demonstrated in this study. The categorization and performance study of wind-catchers as undertaken in this study are the initial steps towards providing a more comprehensive guide of wind tower designs for passive cooling.
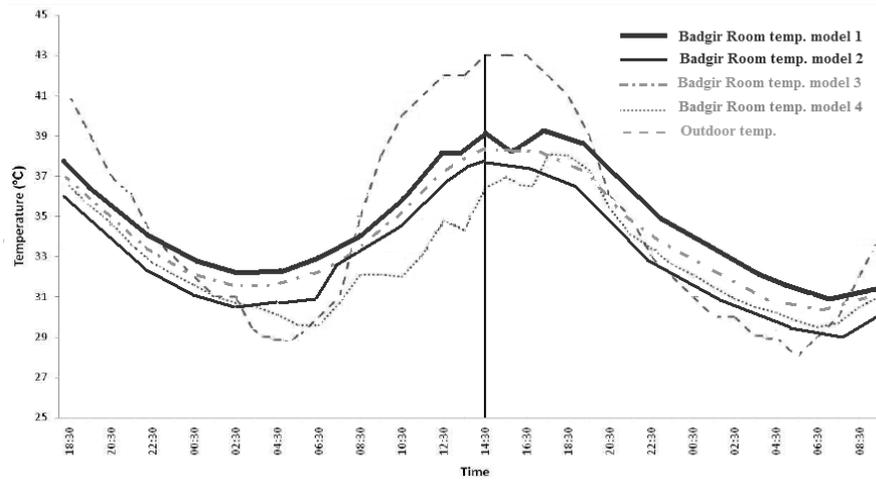
Fig.9   Outdoor and indoor temperature of models
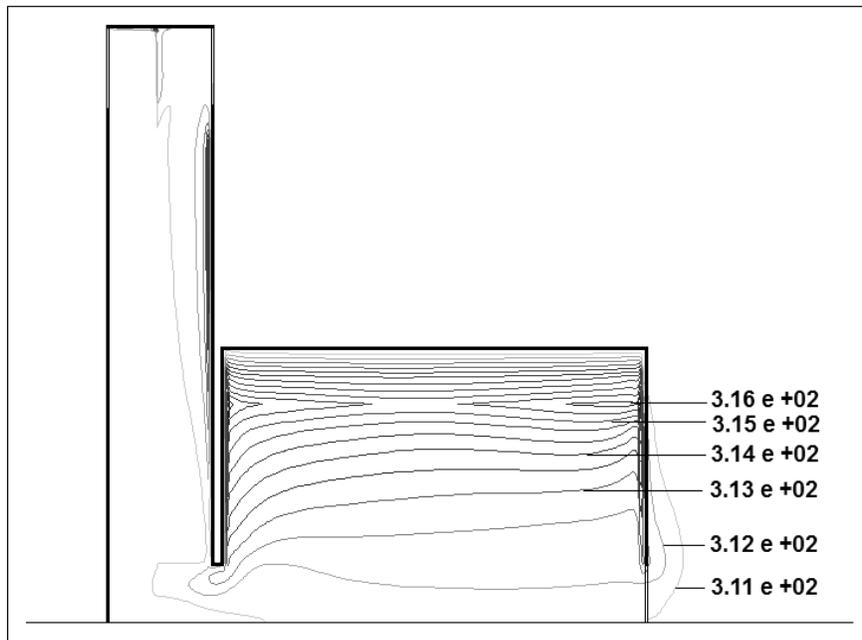
Fig.10   temperature contour of x-blade form wind catcher
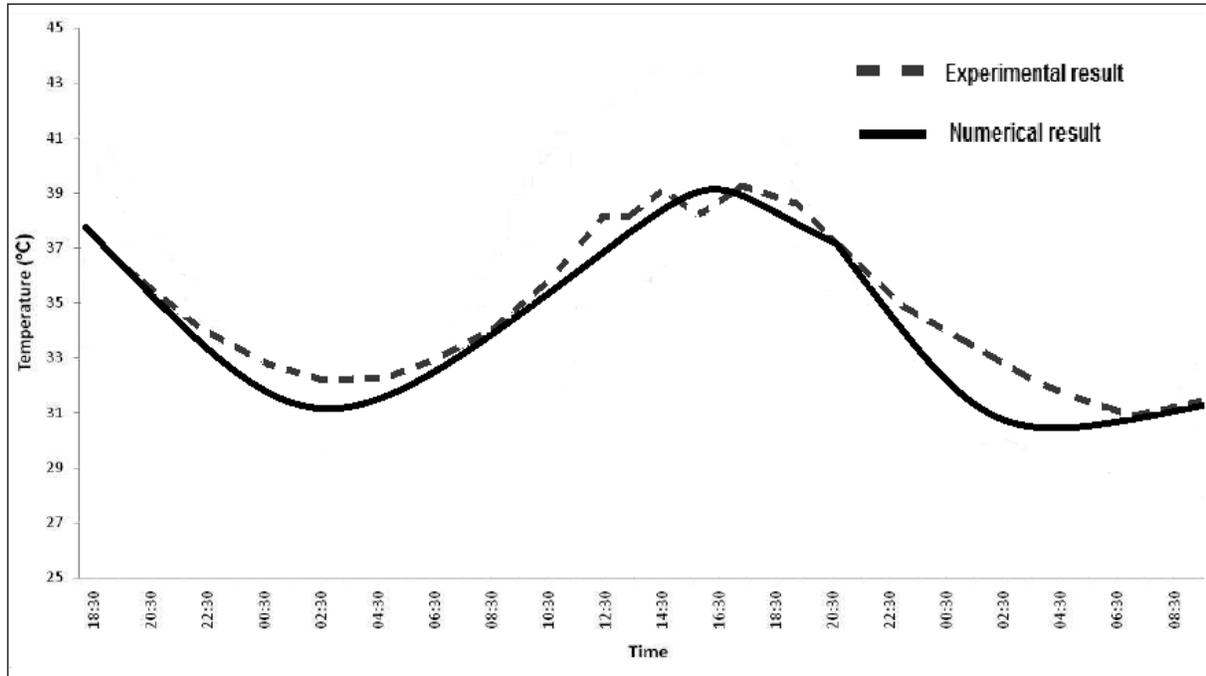
Fig.11   Comparison of numerical and experimental result of x-blade wind catcher

REFERENCES

[1] M.K. Pirnia, Wind catcher, *Journal of Art and Architecture, 10(1),* 1971,30-34

[2] P.Ghaemmaghami, Wind Tower a natural cooling system in Iranian traditional architecture, *1ˢᵗ international conference passive and low energy cooling for the built environment,* Santorini, Greece, 2005, 71-76

[3] A. azami, Badgir in traditional Iranian Architecture, *1ˢᵗ international conference passive and low energy cooling for the built environment,* Santorini, Greece, 2005, 1021-1027

[4] S. Roaf, *The wind catchers of Yazd*, PhD. Thesis, Oxford Polytechnic, 1989

[5] G.H. Memarian, *Residental architecture of Iran* (Soroshe danesh, 1996)

[6] M. Mahmudi, *Badgir, the symbol of Iranian architecture,*(Yazda, 2009)

[7] A. Mahyari, The wind catcher, PhD. Thesis, Sydney university, Australia, 1996

[8] M. Mahmudi, Analysis on Iranian Wind Catcher and its Effect on Natural Ventilation as a solution toward Sustainable Architecture, *World Academy of science, Engineering and technology*, 2009, 574-575.

Performance, Emission and Fuel Induction System of Hydrogen Fuel Operated Spark Ignition Engine - A Review

Shivaprasad K.V[1], Dr. Kumar G.N[2], Dr. Guruprasad K.R[3]

[1]Research Scholar, Department of Mechanical Engineering, National Institute of Technology Karnataka, Surathkal, Srinivasnagar-575025, Mangalore, India.
[2]Assistant Professor, Department of Mechanical Engineering, National Institute of Technology Karnataka, Surathkal, Srinivasnagar-575025, Mangalore, India.
[3]Assistant Professor, Department of Mechanical Engineering, National Institute of Technology Karnataka, Surathkal, Srinivasnagar-575025, Mangalore, India.

ABSTRACT

Fast depletion of fossil fuels and their detrimental effect to the environment is demanding an urgent need of alternative fuels for meeting sustainable energy demand with minimum environmental impact. A lot of research is being carried throughout the world to evaluate the performance, exhaust emission and combustion characteristics of the existing engines using several alternative fuels such as hydrogen, compressed natural gas (CNG), alcohols, liquefied petroleum gas (LPG), biogas, producer gas, bio-diesels, and others. Expert studies indicate hydrogen is one of the most promising energy carriers for the future due to its superior combustion qualities and availability.

This paper provides a comprehensive overview of hydrogen as a fuel for Spark Ignition (SI) internal combustion engine. Discussed topics are introduction to hydrogen, its basic properties, flexibility of hydrogen as a fuel for SI engine, performance and emissions of hydrogen fuel operated SI engine. Also it includes the most significant advances and developments made on the technical adaptations in the SI engine which operate with hydrogen. Finally, it describes the best design of the fuel induction system for SI engines when they are fed with hydrogen.

Keywords – **Autoignition, Combustion, Flammability, Knocking**

1. INTRODUCTION

Hydrogen shows considerable promise as a primary energy carrier in the future. First, it can be produced directly from all primary energy sources, enabling energy feedstock diversity for the transportation sector. These alternative energy resources include wind, solar power, and biomass (plant material), which are all renewable fuel sources. Electricity produced from nuclear fission, or fusion, has also been mentioned with increasing frequency as a possible source of hydrogen (H_2) production through electrolysis of water or thermo chemical cycles. A major benefit of increased H_2 usage for power generation and transportation is that all of these sources minimize our dependence on non-renewable fossil fuels and diversify our energy supply for utilization in end-use energy sectors. Alternatively, H_2 can be produced through coal gasification, or by "steam reforming" of natural gas (NG), both of which are non-renewable fossil fuels but are abundantly available throughout the world. Combining the latter technologies

with carbon capture and storage (CCS) would provide a significant increase in sources same time eliminating green-house gas emissions. Second, since all conventional fossil fuels contain carbon atoms in addition to hydrogen atoms, carbon dioxide (CO_2) is a major product gas formed during the conversion of the fuel to energy. The release of stored chemical energy in H_2 of useful heat produces only water as a product, thus eliminating CO_2, a significant contributor to climate change as a significant greenhouse gas. The concept of H_2 as an energy carrier is often discussed. This concept can best be described in the context of H_2 produced directly from water using electrolysis. While this process is not economically attractive at current costs, if the electricity required to convert H_2O to H_2 is provided by wind or solar power, then the H_2 is produced without creating any CO_2. Given the intermittent nature of wind and solar power sources, the surplus energy produced during the very windy or bright sunny days could be used to produce H_2 that is stored for later use. Under these conditions the stored H_2 becomes an energy carrier that can be used later to produce power where it is needed, either in conjunction with a fuel cell to produce electricity or in a combustor to produce power (internal combustion engine (ICE) or turbine). It should be noted that there are many different ways to produce H_2 from a primary energy feedstock which are usually much more efficient than simple electrolysis, but this serves as a good example of how one might store energy in hydrogen as a carrier for use later. The energy stored in hydrogen can be converted to useful energy through either fuel cells to directly produce electricity or combustion to produce power. However, much more work is needed to improve fuel cells cost effectiveness for everyday use in the general population. For example, fuel cells were recently estimated to be ten times more expensive to produce than a power-comparable ICE.

Another energy conversion technology that is also well proven is combustion. Direct combustion of conventional fossil fuels has been used for centuries and has been refined considerably in recent years due to increasingly stringent pollutant emissions limits and higher-efficiency requirements brought on by recent fuel shortages. As will be described in greater detail in the following sections, recent research indicates that while some technical challenges exist, H_2 can be successfully burned in conventional combustion systems with minimal design changes. In particular, with regard to systems based on dilute premix combustion technology, the unique characteristics of H_2

offer several advantages over conventional hydrocarbon fuels [1].

II. HYDROGEN AS A COMBUSTIBLE FUEL

The use of hydrogen in internal combustion engines may be part of an integrated solution to the problem of depletion of fossil fuels and pollution of the environment. Today, the infrastructure and technological advances in matters of engines can be useful in the insertion of hydrogen as a fuel. Hydrogen has a wide flammability range in comparison with all other fuels. As a result, hydrogen can be combusted in an internal combustion engine over a wide range of fuel-air mixtures. Hydrogen has very low ignition energy. A significant advantage of this is that hydrogen can run on a lean mixture and ensures prompt ignition. Generally, fuel economy is greater and the combustion reaction is more complete when an IC engine runs on a lean mixture. Unfortunately, the low ignition energy means that hot gases and hot spots in the cylinder can serve as sources of ignition, creating problems of premature ignition. Preventing this is one of the challenges associated with running an engine on hydrogen. [2]

The high flame speed of hydrogen, hydrogen engine is much closer to the ideal constant volume combustion than a gasoline engine which produces the reduced exhaust losses and increased engine thermal efficiency [3]. Autoignition temperature of hydrogen is very high. This means that hydrogen is most suitable as a fuel for spark ignition (SI) engines and it is very difficult to ignite hydrogen just by the compression process.

The high diffusion speed of hydrogen also improves the homogeneity of in the cylinder mixture which helps the fuel be completely burnt [4]. Since the flame speed of hydrogen is about five times as large as that of gasoline, retarded spark timing should be adopted for hydrogen engines to ensure its power output and prevent knock [5]-[6]. However, due to the high adiabatic flame temperature of hydrogen, at a specified excess air ratio, NO_x emissions from hydrogen engines are generally much higher than those from gasoline engines, which limit the wide commercialization of the pure hydrogen-fueled engines to some extent [7]-[8]. Besides, the production and storage of hydrogen are still costly at present, which are also the barriers for the pure hydrogen engines to be commercialized in the near future.

The energy density of hydrogen on a mass basis is higher than that of gasoline. However, since hydrogen is very light, the energy density of hydrogen on a volume basis is only 10.8 MJ/m^3, which may lead to a reduced power output for hydrogen engines at stoichiometry conditions, compared with gasoline engines [9].

III. PERFORMANCE CHARACTERISTICS OF HYDROGEN FUELED SI ENGINE

The properties of hydrogen, in particular its wide flammability limits, make it an ideal fuel to combine with other fuels and thereby improve their combustion properties. There are different ways to use hydrogen as a fuel; it can be used as an additive in a hydrocarbon mixture, or as an only fuel, in the presence of air.

The use of the hydrogen as a fuel in the engines has been studied by different authors in the last decade with several degrees of success [10]-[11]-[12]. However, these reports are not necessarily consistent among several researchers. The tendency in this type of reports is focused on the results obtained for specific engines under very narrow operation conditions, and also made emphasis on the emissions and considerations of efficiency. It should be taken into account what has been achieved in this field, focused on the attractive features as in the limitations associated with the disadvantages that are needed to overcome the hydrogen broadly acceptable as a fuel for engines.

Table
Physical Properties of Hydrogen compared to Gasoline and Methane

Property	Hydrogen	Gasoline	Methane
Chemical farmula	H_2	C_8H_{18}	CH_4
Flammability limits (Φ)	0.1–7.1	0.7–4	0.4–1.6
Minimum ignition energy (mJ)	0.02	0.25	0.28
Laminar flame speed at NTP (m/s)	1.90	0.37–0.43	0.38
Adiabatic flame temp (K)	2318	2470	2190
Auto ignition temperature (K)	858	500–750	813
Quenching distance at NTP (mm)	0.64	2.0	2.03
Density at 1 atm and 300 K (kg/m^3)	0.082	730	0.651
Stoichiometric composition in air (% by volume)	29.53	1.65	9.48
Stoichiometric fuel/air mass ratio	0.029	0.0664	0.058
High heating value (MJ/kg)	141.7	48.29	52.68
Low heating value (MJ/kg)	119.7	44.79	46.72
Combustion energy per kg of stoich. mixt. (MJ)	3.37	2.79	2.56
Kinematic viscosity at 300 K (mm^2/s)	110	1.18	17.2
Diffusion coefficient into air at NTP (cm^2/s)	0.61	0.05	0.189

It is also necessary to indicate the practical steps to incorporate the different experimental condition in the existent commercial engines to operate with hydrogen gas. White et al, [13] were made a technical revision of the internal combustion engines operate with hydrogen; their work was an emphasis in the use of hydrogen/gas mixtures with light and heavy load in order to reduce the bad combustion engines. Also, they report the effect of variation in the concentration of the mixture hydrogen/air versus the emissions of NO_x.

M.A Escalante Soberanis, A.M Fernandez also carried out a technical revision on internal combustion engine run with hydrogen fuel. They reported the thermal efficiency of an engine fueled with hydrogen can overcome to that achieved with a gasoline engine (38.9% with hydrogen and 25% with gasoline). The power output of an engine fueled with hydrogen has reached, in laboratory tests, an 80% of that reached by a gasoline engine [14].

Hariganesh R et al, In the Madras Institute of Technology, a comparison study between gasoline and hydrogen as fuels was made [15]. For this purpose, a single cylinder spark ignition engine was adapted to be fueled with hydrogen by injection in the intake manifold. The results of UHC emissions showed that, using hydrogen as a fuel, the levels were near zero, while with gasoline it would maintain over the 2500 rpm, at different requirements of power output. The specific fuel consumption, working with hydrogen, is less than the half than that of gasoline, due to the low energy

density of hydrogen. For the case of nitric oxides emissions, it was reported higher levels in hydrogen combustion. The emissions of the first mixture were about 8000 ppm at an equivalence ratio of 0.85, while for gasoline it was reported 2000 ppm at an equivalence ratio of 1.03, approximately. The minimum ignition energy and the wide range of flammability of hydrogen allow the presence of combustion at lower equivalence ratios than those with gasoline, and it can obtain a higher power at specific equivalence ratios. The high power output of the engine, running with hydrogen, was about 80% of the power reached with gasoline. Hydrogen engine recorded higher volumetric efficiency, compared with that of gasoline, with a power output between 2 and 7 kW, was observed. In the case of thermal efficiency, it reached a maximum of about 27%, at different speeds, over that with gasoline which is about 25%.

S. Verhelst et al, conducted experiments on a Volvo four cylinder sixteen valve gasoline engine with a total swept volume of 1783 cc and a compression ratio of 10.3:1 with some modifications [16]. Their results are presented of the brake thermal efficiency of a bi-fuel hydrogen/gasoline engine, at several engine speeds and loads. They revealed that the efficiencies of both gasoline and hydrogen can be seen to increase as the delivered torque increases. As a result of the increasing torque, the mechanical efficiency increases strongly. For gasoline, the flow losses across the throttle valve increase because of the larger flow, although this is slightly compensated by a larger throttle position (TP). The increase in mechanical efficiency is clearly the dominating effect. In the case of hydrogen, the flow losses have decreased because of a smaller air flow since more air is displaced by hydrogen as a result of the rich mixture. This also leads to a decreased influence of the engine speed on the hydrogen break thermal efficiency (BTEs) too rich with backfire as a consequence. They investigated from their experiment that at low loads (torque outputs of 20Nm) the brake thermal efficiency of hydrogen is (much) higher than on gasoline, the hydrogen BTEs are 40–60% higher relative to the gasoline BTEs. This difference is due to the absence of throttling losses and the lean mixtures of hydrogen. The higher burning velocity of hydrogen is also a contributing factor, as this leads to a more isochoric combustion. The BTE on hydrogen at high speed about 4500 rpm (throttle position for gasoline and $H_2TP = 50\%$, torque output 40 Nm) is about 18% higher relative to gasoline. This difference is not entirely down to a difference in burning velocity however as hydrogen displaces more air due to its low density, throttling losses are lower even though the throttle position is identical, as the lower air flow results in lower flow losses. The gasoline BTE can be seen to be relatively insensitive to the engine speed. For hydrogen, the BTE decreases with engine speed, although the decrease is less pronounced in the wide open throttle (WOT) case. Two effects explain this behavior: first, due to the lean burn operation and large throttle openings, the air flow is much higher in the hydrogen case than for gasoline. This leads to higher flow losses in the intake manifold. For the throttled case, the flow losses include pumping losses so in this case, the BTE is lower and decreases more strongly with engine speed.

Intake-air pressure-boosting (supercharging or turbo charging) is an effective and proven strategy for increasing peak engine power in conventional petroleum-fueled IC engines. For hydrogen engines, pressure-boosting is likely necessary to achieve power densities comparable to petroleum- fueled IC engines. Early work testing boosted hydrogen engine has been carried out by many researchers such as Nagalingam et al [17], Furuhama, Fukuma [18] and Lynch [19]. Nagalingam et al, worked with a single-cylinder research engine and simulated turbocharged operation by pressurizing inlet air to 2.6 bar. In early tests of turbocharged hydrogen engines in commercial vehicles, Lynch converted gasoline and diesel engines to spark-ignited hydrogen operation at maximum inlet pressures of 1.5 bar. Berckmuller et al [20] have reported results from a single-cylinder engine supercharged to 1.8 bar that achieves a 30% increase in specific power output compared to a naturally aspirated gasoline engine. Boosting pressure increases the charge pressure and temperature, the problems of preignition, knock and NO_x control are heightened during boosted operation.

Swain et al, [21] designed an intake manifold to take advantage of the characteristics of hydrogen. The important feature is the use of large passageways with low-pressure drop, which is possible with hydrogen fueling since high intake velocities required for fuel atomization at low engine speeds are not necessary. With the use of a large diameter manifold, Swain et al. reported a 2.6% increase in peak power output compared to that for a small diameter manifold. However, the improvement was lower than the estimated 10% that was expected. One possible explanation for the less than expected performance improvement was that the intake flow dynamics with hydrogen fueling are more complex than for gasoline-fueled engines. In this context, Sierens and Verhelst [16] found that the start and duration of injection influences the volumetric efficiency due to the interaction between the injected hydrogen and the intake pressure waves.

In a technical center of the company Toyota, in Belgium, an experimental research about the combustion characteristics and a spark ignition engine performance, with a four stroke single cylinder, was made [22]. Hydrogen Injection during the intake stroke (300^0 crank angle (CA) before top dead centre (BTDC)) inhibits backfire, but thermal efficiency and power output are limited by knocking and decrease of volumetric efficiency. Hydrogen injection during the compression stroke (130^0 CA BTDC) prevents knocking, increases thermal efficiency and maximizes the power output. However, it was observed that, retarding injection timing during the compression stroke, can lead to a thermal efficiency of about 38.9%, reducing at the same time emissions of nitric oxides.

Changwei Ji, Wang S F [23] carried out the experiments on a modified four-cylinder hybrid hydrogen gasoline engine equipped with an electronically controlled hydrogen port injection system and a hybrid electronic control unit applied to govern the spark timings, injection timings and durations of hydrogen and gasoline. For given hydrogen blending fraction and excess air ratio, the engine load, which was represented by the intake manifolds absolute pressure (MAP), was increased by increasing the opening of the throttle valve. The experimental results demonstrated that the engine brake mean effective pressure (Bmep) was increased after hydrogen addition only at low load conditions. However, at high engine loads, the hybrid hydrogen gasoline engine (HHGE) produced smaller Bmep than the original engine. The engine brake thermal efficiency was distinctly raised with the increase of MAP for

both the original engine and the HHGE. The coefficient of variation in indicated mean effective pressure (COVimep) for the HHGE was reduced with the increase of engine load. The addition of hydrogen was effective in improving gasoline engine operating instability at low load and lean conditions. HC and CO emissions were decreased and NO$_x$ emissions were increased with the increase of engine load. The influence of engine load on CO$_2$ emission was insignificant. All in all, the effect of hydrogen addition to improving engine combustion and emissions performance was more pronounced at lower loads than at high loads.

Changwei Ji and Shuofeng Wang, [24] investigated the idle performance of a spark ignited gasoline engine with hydrogen addition. The research results show that, with the increasing of hydrogen enrichment levels, the engine idle speed remains approximately at its original target. The hydrogen-enriched SI engine gains a higher indicated thermal efficiency and a lower energy flow rate than the pure gasoline SI engine at idle and stoichiometric conditions. The flame development, propagation durations and COVimep are reduced with the increasing hydrogen fraction. Since hydrogen has a wide flammability and fast burning velocity, the CO and HC emissions are reduced with the hydrogen enrichment at idle and lean conditions. Due to the lower peak cylinder temperature, the NO$_x$ emissions are also reduced for the hybrid hydrogen gasoline engine at idle and lean conditions.

Erol Kahraman et al [25] experimentally investigated a conventional four cylinder spark ignition engine operated on hydrogen and gasoline. The compressed hydrogen at 20 MPa has been introduced to the engine adopted to operate on gaseous hydrogen by external mixing. In order to prevent backfire, they were installed the mixer between the carburetor body and inlet manifold at an engine speed above 2600 rpm. Specific features of the use of hydrogen as an engine fuel have been analyzed. The test results have been demonstrated that power loss occurs at low speed hydrogen operation whereas high speed characteristics compete well with the gasoline operation. But, fast burning characteristics of hydrogen permit high speed engine operation. This allows an increase in power output and efficiencies, relatively. NO$_x$ emission of hydrogen fueled engine is about 10 times lower than gasoline fueled engine. The slight traces of CO and HC emissions presented at hydrogen fueled engine are due to the evaporating and burning of lubricating oil film on the cylinder walls. Short time of combustion produces a lower exhaust gas temperature for hydrogen. They also suggested that appropriate changes in the combustion chamber together with a better cooling mechanism would increase the possibility of using hydrogen across a wider operating range.

Fanhua Maa, Yituan He, Jiao Deng Long Jiang et al [26] experimentally investigated the effect of the equivalence ratio (Φ) and ignition advance angle (θ) on idle characteristics of a turbocharged hydrogen fueled SI engine. The experimental data as conducted under various operating conditions including different Φ and θ. It is found that, the ignition advance angle at maximum braking torque (MBT) point decreases gradually with the equivalence ratio increasing from 0.4 to 0.9. Indicated thermal efficiency decreases as Φ increases. Emissions of NO$_x$ increase as Φ increases. When Φ is kept constant, the stated emissions increase as θ increases. During idle conditions of hydrogen fueled engine, a lean mixture with a Φ less than 0.4 is suitable, and the θ should be increased appropriately. The maximum cylinder pressure rises with an increase of Φ and θ. The trend of the maximum rate of pressure rise is similar at different Φ. Only under the conditions of $\Phi = 4$ and $\theta < 10^0$ CA the maximum pressure rise rate remains almost unchanged.

From several practical considerations, hydrogen is safer compared to conventional petroleum fuels. Hydrogen is a low density fuel thus leaking hydrogen rises up very rapidly through the air, thus creating an explosion possibility only to the space immediately above the leak. The ignition energy required to ignite an air fuel mixture depends very much on the equivalence ratio. Hydrogen has an extremely low ignition energy compared to gasoline. Based on the lower flammability limit, hydrogen seems to be superior to gasoline. As far as the quenching distance is concerned, hydrogen combustion which can be initiated with a low energy spark becomes difficult to quench. It can be suggested that appropriate changes in the combustion chamber together with a better cooling mechanism would increase the possibility of using hydrogen across a wider operating range. The hydrogen engine combustion and emissions performance were more evident at lower loads than at high loads.

IV. EMISSION CHARACTERISTICS OF HYDROGEN FUELED SI ENGINE

In recent years, the internal combustion engine powered vehicles have been criticized for their role in environmental pollution through exhaust emissions of mainly the oxides of nitrogen (NO$_x$), carbon monoxide (CO), and unburned hydrocarbons (UBHC). Hydrogen is considered to be clean and efficient alternative fuel among the available. Like electricity, hydrogen is an energy carrier not an energy source. Many scientists have worked both experimentally and analytically with internal combustion engine with hydrogen as fuel. Some of those literatures related to hydrogen are discussed with respect to hydrogen fueled spark ignition engine.

A primary advantage of hydrogen over other fuels is that its only major oxidation product is water vapor. The hydrogen is the most abundant material in the universe and during its combustion with air does not produce significant amounts of carbon monoxide (CO), hydrocarbon (HC), smoke, oxides of sulfur (SO$_x$), leads or other toxic metals, sulfuric acid deposition, ozone and other oxidants, benzene and other carcinogenic compounds, carbon dioxide (CO$_2$), formaldehyde and other greenhouse gases. The only undesirable emission is nitric oxide (NO) and nitrogen dioxide (NO$_2$), which are oxides of nitrogen (NO$_x$) which can collect and avoid their emission to the atmosphere.

H$_2$ICE emissions and control techniques have been thoroughly reviewed by many researchers. Das [27] revealed that ultra-lean combustion (i.e., $\Phi=0:5$), which is sufficiently identical with low temperature combustion, is an effective means for minimizing NO$_x$ emissions in ICEs. He compiled data from various sources for tailpipe emissions with exhaust after-treatment. These data show that NO$_x$ emissions at $\Phi > 0:95$ are near zero with the use of a three-way catalyst) three-way catalyst (TWC).

M A Escalante Soberanis, A M Fernandez [14] also carried out a technical revision on internal combustion engine run with hydrogen fuel. They reported the emissions of air/hydrogen mixtures consist mainly of carbon dioxide and

nitric oxides. In the case of NO_x, higher levels of emissions can be observed, due to the higher temperature and flame velocity of hydrogen compared with other fuels, like gasoline. Emissions of UBHC are the product of lubricant oil heating and the use of oil derivatives for engine cooling. Moreover, in Riverside, University of California, a research about exhaust gas recirculation technique in a hydrogen engine at different fuel flows was carried out [28]. For this purpose, a four cylinder Ford engine was adapted, connecting a line which concludes exhaust gases back to the intake air manifold. Fuel flows from 0.78 to 1.63 kg/h, i.e., Φ from 0.35 to 1.02, and introducing exhaust gases in a phased manner, substituting air. These tests were carried out at an engine speed of 1500 rpm. This technique was shown as an effective method to reduce emissions of nitric oxides to less than 10 ppm, getting a better power output than that with lean mixtures ($\Phi< 0.45$, i.e., 14% Vol. of hydrogen), with thermodynamic efficiencies near 31%.

Erol Kahraman [29] studied the performance and emission characteristics of hydrogen fueled spark ignition engine. The compressed hydrogen at 20 MPa has been introduced to the engine adopted to operate on gaseous hydrogen by external mixing. Two regulators have been used to drop the pressure first to 300 kPa, then to atmospheric pressure. The experiments were carried out on a four-cylinder, four stroke spark ignition engine with carburetor as the fuel induction mechanism. The variations of torque, power, brake thermal efficiency, brake mean effective pressure, exhaust gas temperature, and emissions of NO_x, CO, CO_2, HC, and O_2 versus engine speed are compared for a carbureted SI engine operating on gasoline and hydrogen. He found that NO_x emission of hydrogen fueled engine is about 10 times lower than gasoline fueled engines.

James W Heffel [30] conducted the experiments on hydrogen fueled spark ignition engine to ascertain the effect of exhaust gas recirculation and a standard 3-way catalytic converter had on NO_x emissions and engine performance. All the experiments were conducted at a constant engine speed of 1500 rpm and each experiment used a different fuel flow rate, ranging from 0.78 to 1.63 kg/h. The tests are conducted on a 4-cylinder, 2-liter Ford ZETEC engine specifically designed to run on pure hydrogen using a lean-burn fuel metering strategy. It has a compression ratio of 12:1 and uses a sequential port fuel injection system. From experiments conducted, it can be concluded that if the NO_x emission is not taken into considerations then lean burn strategy can produce more torque than EGR strategy. However, if low NO_x emissions (<10 ppm) are a requirement, the EGR strategy can produce almost 30% more torque than the lean-burn strategy. The results of these experiments demonstrated that using EGR is an effective means to lowering NO_x emissions to less than 1 ppm while also increasing engine output torque.

In addition, Nagalingam [17] reported measurements on a single-cylinder hydrogen engine equipped with a supercharger and an exhaust gas recirculation (EGR) system. The results showing NO_x levels below 100 ppm for equivalence ratios less than 0.4 when operating at supercharged intake pressures of 2.6 bar. Using EGR combined with supercharging and a three-way catalyst (TWC) is shown to significantly increase the power output while limiting tailpipe emissions of oxides of nitrogen (NO_x).

V. HYDROGEN FUEL INDUCTION TECHNIQUES FOR SI ENGINE

The fuel induction technique has been found to be playing a very dominant and sensitive role in determining the characteristics of an IC engine. A unit volume of the stoichiometric hydrogen air mixture has only 85% of the calorific value of the gasoline air mixture. This means that the hydrogen pre mixture spark ignition engine has a smaller maximum power output than the gasoline engine. Therefore the methods to supply hydrogen into an engine and the corresponding design of a hydrogen supply system become one of the key problems to be solved in the research on a hydrogen engine.

The structure of a hydrogen fueled engine is not very different from that of a traditional internal combustion engine but if a gasoline engine without any modification were fueled by hydrogen some problems such as small power output, abnormal combustion (e.g. backfire, pre ignition, high pressure rise rate and even knock) and high NO_x emission would occur. So its fuel supply system and combustion system need suitable modification.

The fuel induction technique for an internal combustion engine can be classified into four categories such as carburetion, inlet port injection, inlet manifold injection, and direct cylinder injection.

Carburetion technique is the oldest technique where the carburetion is done by the use of a gas carburetor.

In an intake port injection system both air and fuel enter the combustion chamber during the intake stroke, but are not premixed in the intake manifold.

Inlet manifold injection is another fuel induction technique where the hydrogen injected into the inlet manifold. This method uses the typical properties of hydrogen fuel to a point of advantage. In this technique, the system is so designed that the intake manifold does not contain any combustible mixture thereby avoiding undesirable combustion phenomena and the air being inducted prior to fuel delivery. It provides a precooling effect and thus avoids preignition sources that could be present on the surface. And also helps to quench or at least to dilute any hot residual combustion products that could be present in the compression space near TDC.

Direct cylinder injection of hydrogen into the combustion chamber does have all the benefits of the late injection as charecterised by manifold injection. In addition, the system permits for fuel delivery after the closing of the intake valve and thus, intrinsically precludes the possibility of backfire.

L M Das [31] concluded in his review paper that late fuel injection is a very promising fuel induction technique which avoids the possibility of backfire which can be effected for both two strokes as well as 4 stroke engines. An appropriate TMI system could be considered for a hydrogen engine in order to ensure smooth operation without any undesirable combustion phenomenon.

LS Guo [32] modified four cylinder hydrogen fueled internal combustion engine for a hydrogen injection with fast response solenoid valves, and its electronic control system. A four cylinder four stroke water cooled gasoline engine with spark ignition is refitted to an in-cylinder injection spark ignition hydrogen fueled engine. Their study to be focused mainly on modification of its hydrogen supply system and combustion system to solve such problems as small power output and abnormal combustion in a hydrogen fueled engine. This study shows that the abnormal

combustion such as a backfire, preignition, high pressure rise rate and knock did not occur and performance of the engine could be improved by means of the hydrogen injection system with fast response solenoid valves. The fast response solenoid valve and its electronic control system possess good switch characteristics and very fast response, thus it could satisfy the working requirements of the injection system. The Intel 7987 chip microprocessor could be applied to control ignition and injection timing optimally so as to improve engine performance. In the hydrogen supply system, the hydrogen injector was always not under a high pressure, so preignition caused by the injector leakage at initial stage of starting the engine was avoided.

Alberto Boretti [33] reported direct injection and jet ignition coupled to the port water injection are used to avoid the occurrence of all abnormal combustion phenomena as well as to control the temperature of gases to turbine in a turbocharged stoichiometric hydrogen engine. Port water injection coupled with direct injection and jet ignition may permit the stoichiometric operation of hydrogen engines. This brings the advantages of high power densities, even if at the expenses of reduced peak fuel conversion efficiencies. High pressure and high flow rate direct injection performed close to the ignition top dead center eliminates the occurrence of backfire and contribute to reduce the probability of knock reducing the time in between an ignitable mixture is made available within the cylinder and the time the mixture is fully burned. Top dead center jet ignition produces very fast combustion rates, for an almost isochoric combustion process, with multiple jet of hot reacting gases igniting the mixture in the bulk in multiple locations thus permitting better conversion efficiencies and reduced likeliness of knock occurrence for the reduce time to complete combustion. Port water injection further reduces the likeliness of knock occurrence strongly reducing the charge temperature through vaporization.

Ali Mohammadi et al [22] developed a direct-injection spark ignition hydrogen engine and attention was paid on the effects of injection timing on the engine performance, combustion characteristics and NO_x emission under a wide range of engine loads. From this research it can be revealed that In-cylinder injection of hydrogen during the intake stroke as well as compression stroke prevents backfire and knock respectively. The experiments results suggested that Hydrogen injection at later stage of compression stroke can achieve the thermal efficiency higher than 38.9% and the brake mean effective pressure 0.95MPa and also under high engine output conditions, late injection of hydrogen offers a great reduction in NO_x emission due to the lean operation.

As indicated in the previous researches, employing direct-injection technology in a hydrogen engine is very effective to control the abnormal combustion of hydrogen and achieve high thermal efficiency and output power. Although late injection results in lower NO_x emissions, utilization of other techniques such as exhaust gas recirculation and after treatment methods are required to bring the NO_x emission to acceptable levels.

VI. CONCLUSION

The use of hydrogen in internal combustion engines may be part of an integrated solution to the problem of depletion of fossil fuels and pollution of the environment. Today, the infrastructure and technological advances in matters of engines can be useful in the insertion of hydrogen as a fuel.

There are good prospects for increased efficiencies, high power density, and reduced emissions with hybridization, multi-mode operating strategies, and advancements in ICE design and materials.

The hydrogen infrastructure at the time is not in place to supply hydrogen demands, but with more development using hydrogen as a fuel will motivate the development of the infrastructure.

REFERENCES

[1] Robert W. Schefer, Christopher White and Jay Keller, *Lean Hydrogen Combustion* Elsevier, 2008

[2] Heywood J.B*., Internal Combustion Engine Fundamentals* McGraw-Hill Book Co., 1988.

[3] Shudo T, Nabetani S and Nakajima Y, Analysis of the degree of constant volume and cooling loss in a spark ignition engine fuel led with hydrogen. *International Journal of Engine Research, 2(1)*, 2001, 81–92.

[4] Ma F, Wang J, Wang Y, Wang Y.F, Zhong Z and Ding S, An investigation of optimum control of a spark ignition engine fueled by NG and hydrogen mixtures, *International Journal of Hydrogen Energy, 33(8)*, 2008, 7592–7606.

[5] Salimi F, Shamekhi H.S and Pourkhesalian A.M, Role of mixture richness, spark and valve timing in hydrogen-fueled engine performance and emission. *International Journal of Hydrogen Energy, 34(9)*, 2009, 3922–3929.

[6] Li H and Karim G, Knock in hydrogen spark ignition engines, *International Journal of Hydrogen Energy, 29(8)*, 2004, 859–865.

[7] Ren J.Y, Qin W, Egolfopoulos F.N and Tsotsis T.T, Strain-rate effects on hydrogen-enhanced lean premixed combustion. *International Journal of Combustion and Flame*, *124(4)*, 2001, 717–720.

[8] Haroun Abdul-Kadim Shahad Al-Janabi and Maher Abdul-Resul Sadiq Al-Baghdadi, A prediction study of the effect of hydrogen blending on the performance and pollutants emission of a four stroke spark ignition engine, *International Journal of Hydrogen Energy 24(4)*, 1999, 363–375.

[9] Ji C and Wang S, Experimental study on combustion and emissions performance of a hybrid hydrogen-gasoline engine at lean burn limits, *International Journal of Hydrogen Energy*, *35(3)*, 2010, 1453–1462.

[10] Gopal G, Srinivasa R.P, Gopalakrishnan K.V and Murthy B S, Use of hydrogen in dual-fuel engines, *International Journal of Hydrogen Energy, 7(3)*, 1982, 267-272.

[11] Mathur H.B and Khajuria P.R, Performance and emission characteristics of hydrogen fuelled spark ignition engine, *International Journal of Hydrogen Energy, 9(8)*, 1984, 729-735.

[12] Karim G.A, Hydrogen as a spark ignition engine fuel, *International Journal of Hydrogen Energy*, *28(5)*, 2003, 569-575.

[13] White C.M, Steeper R.R and Lutz A.E, The hydrogen-fueled internal combustion engine: a technical review, *International Journal of Hydrogen Energy, 31(10)*, 2006, 1292–1305.

[14] M.A. Escalante Soberanis and A.M. Fernandez, A review on the technical adaptations for internal combustion engines to operate with gas/hydrogen

mixtures, *International Journal of Hydrogen Energy, 35(21)*, 2010, 12134–12140.

[15] Hari Ganesh R, Subramanian V, Balasubramanian V, Mallikarjuna J.M, Ramesh A and Sharma R.P, Hydrogen fueled spark ignition engine with electronically controlled manifold injection: an experimental study, *Journal of Renewable Energy, 33(6)*, 2008, 1324- 1333.

[16] S Verhelst, P. Maesschalck, N Rombaut and R Sierens, Efficiency comparison between hydrogen and gasoline on a bi-fuel hydrogen/gasoline engine, *International Journal of Hydrogen Energy, 34 (5)*, 2009, 2504-2510.

[17] Nagalingam B, Dubel M, and Schmillen K, Performance of the supercharged spark ignition hydrogen engine, *SAE International, 10(31)*,1983.

[18] Furuhama S and Fukuma T, High output power hydrogen engine with high pressure fuel injection- Hot surface ignition and turbo charging, *International Journal of Hydrogen Energy, 11(6)*, 1986, 399-407.

[19] Lynch F E, Parallel Induction: a simple fuel control method for hydrogen engines, *International Journal of Hydrogen Energy, 8(9)*, 1983, 721-730.

[20] Berckmuller M, Rottengruber H, Eder A, Brehm N, Elsasser G and Muller-Alander G, Potentials of a charged SI-hydrogen engine. *SAE International, 10(17)*, 2003.

[21] Swain, M.R, Schade and G.J Swain, Design and Testing of a Dedicated Hydrogen-Fueled Engine, *SAE International, 5(1)*, 1996.

[22] Ali Mohammadi, Masahiro Shioji, Yasuyuki Nakai Wataru, Ishikura and Eizo Tabo, Performance and combustion characteristics of a direct injection SI hydrogen engine, *International Journal of Hydrogen Energy, 32(2)*, 2007, 296-304.

[23] Changwei Ji C.W and Wang S.F, Effect of hydrogen addition on the idle performance of a spark ignited gasoline engine at stoichiometric condition, *International Journal of hydrogen Energy, 34(8)*, 2009, 3546-3556.

[24] Changwei Ji, Shuofeng Wang and Bo Zhangn., Combustion and emissions characteristics of a hybrid Hydrogen- gasoline engine under various loads and lean conditions, *International Journal of Hydrogen Energy, 35(11)*, 2010, 5714-5722.

[25] Erol Kahraman, S. Cihangir Ozcanl and Baris Ozerdem., Experimental study on performance and emission characteristics of hydrogen fuelled spark ignition engine, *International Journal of Hydrogen Energy, 32(12)*, 2007, 2066 – 2072.

[26] Fanhua Ma, Yituan He, Jiao Deng, Long Jiang, Nashay Naeve,Mingyue Wanga, and Renzhe Chen, Idle characteristics of a hydrogen fueled SI engine, *International journal of hydrogen energy 36(7)*, 2011, 4454- 4460.

[27] L.M. Das, Exhaust emission characterization of hydrogen operated engine system: Nature of pollutants and their control techniques, *International Journal of Hydrogen Energy, 16(11)*, 1991, 765–775.

[28] Heffel J, NO_x emission and performance data for hydrogen fueled internal combustion engine at 1500 rpm using exhaust gas recirculation, *International Journal of Hydrogen Energy,28(8)*, 2003, 901-908.

[29] Erol Kahramana, S Cihangir Ozcanlıb, and Baris Ozerdemb., An Experimental study on performance and emission characteristics of a hydrogen fuelled spark ignition engine, *International Journal of Hydrogen Energy*, 32, 2007, 2066 – 2072.

[30] James W Heffel, NO_x emission reduction in a hydrogen fueled internal combustion engine at 3000 rpm using exhaust gas recirculation, *International Journal of Hydrogen Energy, 28(11)*, 2003, 1285 – 1292.

[31] L M Das, Fuel induction techniques for a hydrogen operated engine, *International Journal of Hydrogen Energy, 15(11)*, 1990, 833-842.

[32] L S Guo, H B Lu and J D Li, A hydrogen injection system with solenoid valves for a four Cylinder hydrogen fuelled engine, *International Journal of Hydrogen Energy, 24(4)*, 1999, 377-382.

[33] Alberto Boretti, Stoichiometric H_2ICEs with water injection, *International Journal of Hydrogen Energy*, *36(7)*, 2011, 4469-4473.

A Novel Approach for TCSC-Based Supplementary Damping Controller Design Using Multi-Objective Optimization Technique

A.K.Baliarsingh, D.P.Dash
Department of Electrical Engineering
Orissa Engineering College, Bhubaneswar, India

S.Panda, B.N.Mohanty
Department of Electrical and Electronics Engineering
VSSUT, Burla, NMIET Bhubaneswar, India

Abstract
Design of an optimal controller requires optimization of multiple performance measures that are often noncommensurable and competing with each other. Design of such a controller is indeed a multi-objective optimization problem. Being a population based approach; Genetic Algorithm (GA) is well suited to solve multi-objective optimization problems. This paper investigates the application of GA-based multi-objective optimization technique for the design of a Thyristor Controlled Series Compensator (TCSC)-based supplementary damping controller. The design objective is to improve the power system stability with minimum control effort. The proposed technique is applied to generate Pareto set of global optimal solutions to the given multi-objective optimization problem. Further, a fuzzy-based membership value assignment method is employed to choose the best compromise solution from the obtained Pareto solution set. Simulation results are presented to show the effectiveness and robustness of the proposed approach.

Keywords-*multi-objective optimization, genetic algorithm, pareto solution set, thyristor controlled series compensator, power system stability*

I.   INTRODUCTION

Real world problems often have multiple conflicting objectives competing with each other. For example, while designing a control system, we would usually like to have a high-performance controller, but we also want to achieve desired performance with little control efforts (cost). Optimization of multiple performance measures which are noncommensurable and competing with each other is in reality a multi-objective optimization problem. In multi-objective optimization problems generally there is no single solution that is the best when measured on all objectives. Hence several trade-off solutions (called the *Pareto optimal set*) are usually preferred [1]. Control systems optimization problems involving the optimization of multiple objective functions require high computational time and effort [2, 3]. As conventional techniques are difficult to apply, modern population based heuristic optimization techniques are preferred to obtain Pareto optimal set [4].

Recent development of power electronics introduces the use of Flexible AC Transmission Systems (FACTS) controllers in power systems [5]. Thyristor Controlled Series Compensator (TCSC) is one of the important members of FACTS family that is increasingly applied with long transmission lines by the utilities in modern power systems [6-11]. The majority of the control methodologies presented in literature employ single objective optimization technique to get the desired performance. This paper proposes to use a multi-objective optimization technique for the optimal TCSC-based controller design.

There are two general approaches to multiple objective optimizations. One approach to solve multi-objective optimization problems is by combining the multiple objectives into a scalar cost function, ultimately making the problem single-objective prior to optimization. However, in practice, it can be very difficult to precisely and accurately select these weights as small perturbations in the weights can lead to very different solutions. Further, if the final solution found cannot be accepted as a good compromise, new runs of the optimiser on modified objective function using different weights may be needed, until a suitable solution is found. These methods also have the disadvantage of requiring new runs of the optimizer every time the preferences or weights of the objectives in the multi-objective function change [4]. The second general approach is to determine an entire Pareto optimal solution set or a representative subset. Pareto optimal solution sets are often preferred to single solutions because they can be practical when considering real-life problems, since the final solution of the decision maker is always a trade-off between crucial parameters [12].

In this paper, the design problem of a TCSC is formulated as a multi-objective optimization problem. Genetic Algorithm- based multi-objective optimization method is adapted for generating Pareto solutions in designing a TCSC-based controller. The design objective is to get maximum damping (performance) with minimum control effort (cost). Further a fuzzy based membership function value assignment method is employed to choose the best compromise solution from the obtained Pareto set. Simulation results are presented under various loading conditions and disturbances to show the effectiveness and robustness of the proposed approach.

II.   MODELING THE POWER SYSTEM WITH TCSC

The single-machine infinite-bus (SMIB) power system installed with a TCSC, shown in Figure 1 is considered in this study. In the Figure, X_T and X_L represent the reactance of the transformer and the transmission line respectively; V_T and V_B are the generator terminal and infinite bus voltage respectively.

In the design of electromechanical mode damping stabilizer, a linearized incremental model around an operating point is usually employed. The Phillips-Heffron model of the power system with FACTS devices is obtained by linearizing nonlinear equations of the power system around an operating condition [12].
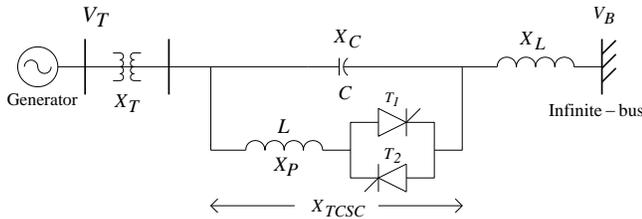

Figure 1. Single machine infinite bus power system with TCSC
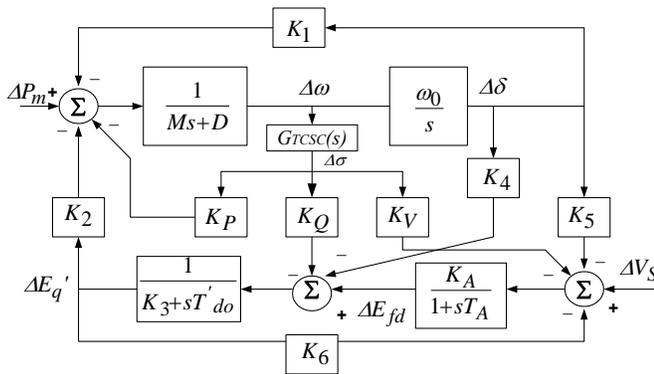
Figure 2. Modified Phillips-Heffron model of SMIB with TCSC

The linearized expressions are as follows [10]:

$$\dot{\Delta\delta} = \omega_b \Delta\omega$$

$$\dot{\Delta\omega} = [-K_1\Delta\delta - K_2\Delta E_q' - K_P\Delta\sigma - D\Delta\omega]/M$$

$$\dot{\Delta E_q'} = [-K_3\Delta E_q' - K_4\Delta\delta - K_Q\Delta\sigma + \Delta E_{fd}]/T_{d0}'$$

$$\dot{\Delta E_{fd}} = [-K_A(K_5\Delta\delta + K_6\Delta E_q' + K_V\Delta\sigma) - \Delta E_{fd}]/T_A$$

Where,

$$K_1 = \partial P_e / \partial \delta, \ K_2 = \partial P_e / \partial E'_q, \ K_P = \partial P_e / \partial \sigma$$

$$K_3 = \partial E_q / \partial E'_q, \ K_4 = \partial E_q / \partial \delta, \ K_Q = \partial E_q / \partial \sigma$$

$$K_5 = \partial V_T / \partial \delta, \ K_6 = \partial V_T / \partial E'_q, \ K_V = \partial V_T / \partial \sigma$$

(1)

The notations in equation (1) for the variables and parameters described are standard and defined in the nomenclature. For more details, the readers are suggested to refer [1, 12]. The Phillips-Heffron model of the SMIB system with TCSC is obtained using the linearized equations. The corresponding block diagram model is shown in Figure 2.

III. PROBLEM FORMULATION

A. TCSC Controller Structure

The commonly used lead–lag structure is chosen in this study as a TCSC controller. The structure of the TCSC controller is shown in Figure 3. It consists of a gain block with gain K_T, a signal washout block and two-stage phase compensation block. The phase compensation block provides the appropriate phase-lead characteristics to compensate for the phase lag between input and the output signals. The signal washout block serves as a high-pass filter, with the time constant T_{WT}, high enough to allow signals associated with oscillations in input signal to pass unchanged. Without it steady changes in input would modify the output. From the viewpoint of the washout function, the value of T_{WT} is not critical and may be in the range of 1 to 20 seconds [13].
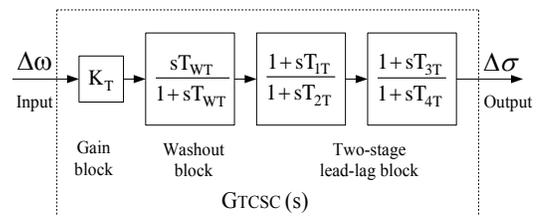

Figure 3. Structure of TCSC-based Controller

B. Objective Function

It is worth mentioning that the proposed controller is designed to damp power system oscillations with minimum control effort following a disturbance. The tuning of lead-lag controller is done, by optimizing the error signal and the control signal values simultaneously. The objective is formulated as the minimization of function F given by:

$$F = (F_1, F_2) \tag{2}$$

Where, $\quad F_1 = \int_0^{t_1} e^2(t) \, dt$ and $F_2 = \int_0^{t_1} u^2(t) \, dt$

In the above equations, 'e' is the error signal i.e. changes in the speed deviation and 'u' is the TCSC control out put i.e. changes in the conduction angle of the TCSC controller and t_1 is the time range of the simulation. For the objective function calculation, the time-domain simulation of the power system model is carried out for the simulation period.

IV. MULTI-OBJECTIVE OPTIMIZATION

A Multi-objective Optimization Problem (MOP) differs from a single-objective optimization problem because it contains several objectives that require optimization. In case of single objective optimization problems, the best single design solution is the goal. But for multi-objective problems, with several and possibly conflicting objectives, there is usually no single optimal solution. Therefore, the decision maker is required to select a solution from a finite set by

making compromises. A suitable solution should provide for acceptable performance over all objectives [14-15].

A general formulation of a MOP consists of a number of objectives with a number of inequality and equality constraints. Mathematically, the problem can be written as [16]:

$$\text{minimize/maximize } f_i(\mathbf{x}) \quad \text{for } i = 1, 2, ..., n. \qquad (3)$$

Subject to constraints:

$$g_j(\mathbf{x}) \leq 0 \qquad j = 1, 2, ..., J$$

$$h_k(\mathbf{x}) \leq 0 \qquad k = 1, 2, ..., K$$

where

$$f_i(\mathbf{x}) = \{ f_1(\mathbf{x}), ... f_n(\mathbf{x}) \}$$

n = number of objectives or criteria to be optimized

$\mathbf{x} = \{x_1, ..., x_p\}$ is a vector of decision variables

p = number of decision variables

There are two approaches to solve the MOP. One approach is the classical weighted-sum approach where the objective function is formulated as a weighted sum of the objectives. But the problem lies in the correct selection of the weights or utility functions to characterise the decision-makers preferences. In order to solve this problem, the second approach called Pareto-optimal solution can be adapted. The MOPs usually have no unique or perfect solution, but a set of non-dominated, alternative solutions, known as the Pareto-optimal set. Assuming a minimisation problem, dominance is defined as follows:

A vector $\mathbf{u}=(u_1,....,u_n)$ is said to dominate $\mathbf{v}=(v_1,.....,v_n)$ if and only if $\mathbf{u}$ is partially less than $\mathbf{v}$ ($\mathbf{u}$ p< $\mathbf{v}$),

$$\forall \; i \, \varepsilon \{1,...,n\}, u_i \leq v_i \wedge \exists \; i \, \varepsilon \{1,...,n; \; u_i < v_i \qquad (4)$$

A solution $x_u \, \varepsilon \, U$ is said to be Pareto-optimal if and only if there is no $x_v \, \varepsilon \, U$ for which $\mathbf{v} = \mathbf{f}(x_v) = (v_1,...,v_n)$ dominates $\mathbf{u} = f(x_u) = (u_1,...,u_n)$.

The ability to handle complex problems, involving features such as discontinuities, multimodality, disjoint feasible spaces and noisy function evaluations reinforces the potential effectiveness of GA in optimization problems. Although, the conventional GA is also suited for some kinds of multi-objective optimization problems, it is still difficult to solve those multi-objective optimization problems in which the individual objective functions are in the conflict condition.

Being a population based approach; GA is well suited to solve MOPs. A generic single-objective GA can be easily modified to find a set of multiple non-dominated solutions in a single run. The ability of GA to simultaneously search different regions of a solution space makes it possible to find a diverse set of solutions for difficult problems with non-convex, discontinuous, and multi-modal solutions spaces. The crossover operator of GA exploits structures of good

solutions with respect to different objectives to create new non-dominated solutions in unexplored parts of the Pareto front. In addition, most multi-objective approach does not require the user to prioritise, scale, or weigh objectives. In this paper, real-coded genetic algorithm (RCGA) optimization technique has been used to solve the given MOP problem. A brief overview of RCGA has been provided in the next section.

V. REAL CODED GENETIC ALGORITHM

Recently, Genetic Algorithm (GA) appeared as a promising evolutionary technique for handling the optimization problems [17]. GA has been popular in academia and the industry mainly because of its intuitiveness, ease of implementation, and the ability to effectively solve highly nonlinear, mixed integer optimisation problems that are typical of complex engineering systems. It has been reported in the literature that Real-Coded Genetic Algorithm (RCGA) is more efficient in terms of CPU time and offers higher precision with more consistent results. Implementation of GA requires the determination of six fundamental issues: chromosome representation, selection function, the genetic operators, initialization, termination and evaluation function. Brief descriptions about these issues are provided in the following sections.

A. Chromosome representation

Chromosome representation scheme determines how the problem is structured in the GA and also determines the genetic operators that are used. Each individual or chromosome is made up of a sequence of genes. Various types of representations of an individual or chromosome are: binary digits, floating point numbers, integers, real values, matrices, etc. Generally natural representations are more efficient and produce better solutions. Real-coded representation is more efficient in terms of CPU time and offers higher precision with more consistent results.

B. Selection function

To produce successive generations, selection of individuals plays a very significant role in a genetic algorithm. The selection function determines which of the individuals will survive and move on to the next generation. A probabilistic selection is performed based upon the individual's fitness such that the superior individuals have more chance of being selected. There are several schemes for the selection process: roulette wheel selection and its extensions, scaling techniques, tournament, normal geometric, elitist models and ranking methods.

The selection approach assigns a probability of selection P_j to each individuals based on its fitness value. In the present study, normalized geometric selection function has been used. In normalized geometric ranking, the probability of selecting an individual P_i is defined as:

$$Pi = q' (1-q)^{r-1} \qquad (5)$$

$$q' = \frac{q}{1 - (1-q)^P} \qquad (6)$$

Where,

q = probability of selecting the best individual

r = rank of the individual (with best equals 1)

P = population size

C. Genetic Operators

The basic search mechanism of the GA is provided by the genetic operators. There are two basic types of operators: crossover and mutation. These operators are used to produce new solutions based on existing solutions in the population. Crossover takes two individuals to be parents and produces two new individuals while mutation alters one individual to produce a single new solution. The following genetic operators are usually employed: simple crossover, arithmetic crossover and heuristic crossover as crossover operator and uniform mutation, non-uniform mutation, multi-non-uniform mutation, boundary mutation as mutation operator. Arithmetic crossover and non-uniform mutation are employed in the present study as genetic operators. Crossover generates a random number r from a uniform distribution from 1 to m and creates two new individuals by using equations:

$$x_i' = \begin{cases} x_i, & if\ i < r \\ y_i\ otherwise \end{cases} \quad (7)$$

$$y_i' = \begin{cases} y_i, & if\ i < r \\ x_i\ otherwise \end{cases} \quad (8)$$

Arithmetic crossover produces two complimentary linear combinations of the parents, where r = U (0, 1):

$$\bar{X}' = r\,\bar{X} + (1-r)\,\bar{Y} \quad (9)$$

$$\bar{Y}' = r\,\bar{Y} + (1-r)\,\bar{X} \quad (10)$$

Non-uniform mutation randomly selects one variable j and sets it equal to a non-uniform random number.

$$x_i' = \begin{cases} x_i + (b_i - x_i)\,f(G) & if\ r_1 < 0.5, \\ x_i + (x_i + a_i)\,f(G) & if\ r_1 \geq 0.5, \\ x_i, & otherwise \end{cases} \quad (11)$$

Where,

$$f(G) = (r_2 (1 - \frac{G}{G_{max}}))^b \quad (12)$$

r_1, r_2 = uniform random nos. between 0 to 1.

G = current generation.

G_{max} = maximum no. of generations.

b = shape parameter.

D. Initialization, termination and evaluation function

An initial population is needed to start the genetic algorithm procedure. The initial population can be randomly generated or can be taken from other methods. GA moves from generation to generation until a stopping criterion is met. The stopping criterion could be maximum number of generations, population convergence criteria, lack of improvement in the best solution over a specified number of generations or target value for the objective function. Evaluation functions or objective functions of many forms can be used in a GA so that the function can map the population into a partially ordered set.


Figure 4. Pareto optimal solution surface

VI.   RESULTS AND DISCUSSIONS

The objective function given by equation (2) is evaluated by simulating the system dynamic model considering a 10 % step increase in mechanical power input (ΔP_m) at t = 1.0 sec. For the implementation of RCGA normal geometric selection is employed which is a ranking selection function based on the normalized geometric distribution. Arithmetic crossover takes two parents and performs an interpolation along the line formed by the two parents. Non uniform mutation changes one of the parameters of the parent based on a non-uniform probability distribution. This Gaussian distribution starts wide, and narrows to a point distribution as the current generation approaches the maximum generation. Using the above approach the final Pareto solution surface is obtained as shown in Figure 4 where the Pareto solutions are shown with the marker 'o'.

A. Best Compromise Solution

In the present paper, a Fuzzy-based approach is applied to select the best compromise solution from the obtained Pareto set. The j-th objective function of a solution in a Pareto set f_j is represented by a membership function μ_j defined as [18]:

$$\mu_j = \begin{cases} 1, & f_j \leq f_j^{\min} \\ \dfrac{f_j^{\max} - f_j}{f_j^{\max} - f_j^{\min}}, & f_j^{\min} < f_j < f_j^{\max} \\ 0, & f_j \geq f_j^{\max} \end{cases} \qquad (13)$$

Where $f_j^{\max}$ and $f_j^{\min}$ are the maximum and minimum values of the j-th objective function, respectively.

For each solution i, the membership function μ^i is calculated as:

$$\mu^i = \frac{\displaystyle\sum_{j=1}^{n} \mu_j^i}{\displaystyle\sum_{i=1}^{m} \sum_{j=1}^{n} \mu_j^i} \qquad (14)$$

Where, n is the number of objectives functions and m is the number of solutions. The solution having the maximum value of μ^i is the best compromise solution.

Using the above approach the best compromise solution is obtained as:

$K_T = 38.278$, $T_{1T} = 0.5632s$, $T_{2T} = 0.2646s$, $T_{3T} = 0.1013$ s and $T_{4T} = 0.1549$ s.

B. Simultion Results

In order to verify the effectiveness of the proposed approach, the performance of the proposed TCSC controller is tested for different loading conditions. The mechanical power input to the generator is increased by 5 % at t =1.0 sec at nominal loading condition (Pe=0.9 pu). The system response for the above contingency is shown in Figures 5 and 6. For comparison, Figures 5 and 6 show the response when the best compromise solution is used (shown in the Figures in solid lines) and also when the other two solutions from the Pareto set are used (shown in dotted and dashed lines). It can be seen from Figures 5 and 6 that when both the
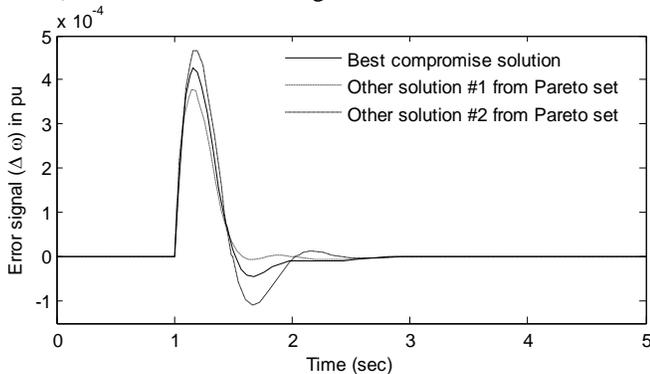
Figure 5. System error response for disturbance in P_m
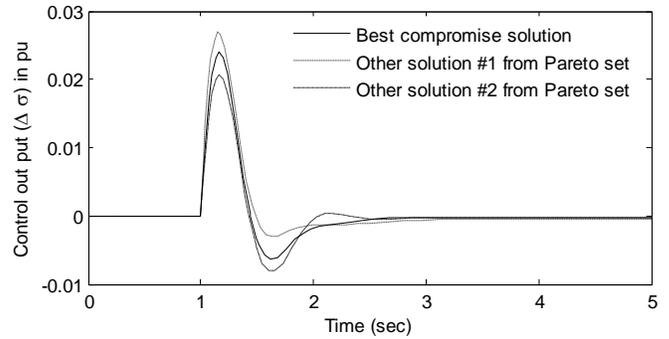
Figure 6. System control output response for disturbance in P_m
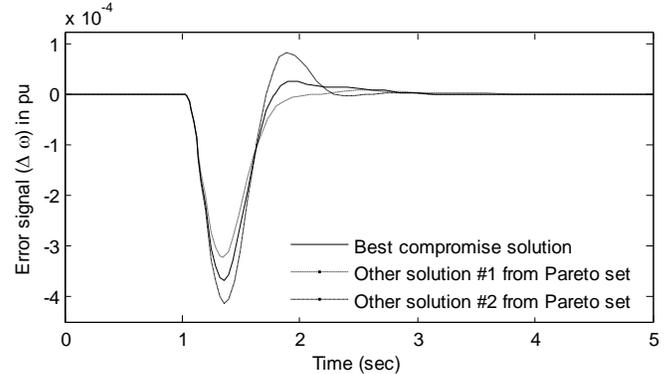
Figure 7. System error response for disturbance in V_{ref}
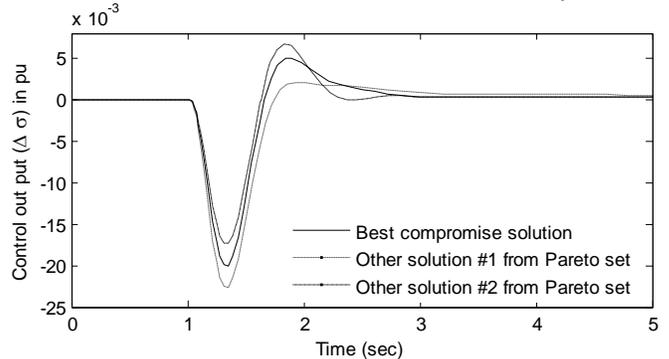
Figure 8. System control output response for disturbance in V_{ref}

error and control out put is considered, the proposed best compromise is the best choice.

To test the robustness of the proposed controller the loading condition is changed to light loading condition (Pe=0.4 pu) and a 5 % step increase in reference voltage is considered at t = 1.0 sec. The system responses for the above cases are shown in Figures 7-8. It can be seen from these Figures that proposed controller is robust and works effectively under various operating conditions and disturbances. Also it can be seen that when both the objectives are considered the proposed approach gives a better response.

VII. CONCLUSIONS

In this study optimal design of a TCSC-based controller is presented and discussed. The design objective is to improve the stability of a power system with minimum control effort. A real-coded genetic algorithm based solution technique is applied to generate a Pareto set of global optimal solutions to the given multi-objective optimization problem. Further, a fuzzy-based membership value assignment method is employed to choose the best compromise solution from the obtained Pareto solution set. Simulation results are presented at various loading conditions and disturbances to show the effectiveness and robustness of the proposed approach.

The proposed method is valuable for the design of the interactive decision making. The decision makers can choose from the solutions in the Pareto-optimal set to find out the best solution according to the requirement and needs as the desired parameters of their controllers. The results show that evolutionary algorithms are effective tools for handling multi-objective optimization where multiple Pareto-optimal solutions can be found in one simulation run.

APPENDIX

Static System data: All data are in pu unless specified otherwise.

Generator:

$M = 9.26$ *s.*, $D = 0$, $X_d = 0.973$, $X_q = 0.55$,

$X'_d = 0.19$, $T'_{do} = 7.76$, $f = 60$, $V_T = 1.05$,

$X_{TL} + X_T = 0.997$, *Excitor:* $K_A = 50$, $T_A = 0.05$ *s*

TCSC Controller: $X_{TCSC0} = 0.2169$,

REFERENCES

[1] P. Kundur, Power System Stability and Control. New York: McGraw-Hill, 1994.
[2] K. Deb, Multi-Objective Optimization using Evolutionary Algorithms, Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons, 2001.
[3] J.R.H. Carvalho, P.A.V. Ferreira, "Multiple-criterion control: A convex programming approach", Automatica, vol. 31, pp. 1025-1029, 1995.
[4] L.Z. Liao, D. Li, "Adaptive differential dynamic programming for multi-objective optimal control", Automatica, vol. 38, pp. 1003-1015, 2002.
[5] C.A.C. Coello, "A comprehensive survey of evolutionary-based multi-objective optimization", Knowledge and Information Systems, vol. 1, no. 3, pp. 269-308, 1999.
[6] N.G. Hingorani, L. Gyugyi, Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems, IEEE Press, New York, 2000.
[7] Y. H Song, T. A. Johns, Flexible AC Transmission Systems (FACTS), IEE, London, 2000.
[8] R. M Mathur, R. K. Verma, Thyristor-based FACTS Controllers for Electrical Transmission Systems, IEEE Press, Piscataway, 2002.
[9] A. D Del Rosso, C. A Canizares and V.M. Dona, "A study of TCSC controller design for power system stability improvement," IEEE Trans. Power Systs., vol. 18, pp. 1487-1496. 2003.
[10] S. Panda, N. P. Padhy, R. N. Patel, "Modeling, simulation and optimal tuning of TCSC controller", International Journal of Simulation Modelling. vol. 6, no. 1, pp. 7-48, 2007.
[11] S. Panda, and N. P. Padhy, "Comparison of particle swarm optimization and genetic algorithm for FACTS-based controller design", Applied Soft Computing. vol. 8, pp. 1418-1427, 2008.
[12] V. Chankong, and Y. Haimes, Multiobjective Decision Making Theory and Methodology, New York: North-Holland, 1983.
[13] K.R. Padiyar, Power System Dynamic Stability and Control. BS Publications, 2nd Edition, Hyderabad, India, 2002.
[14] S. Das, B.K. Panigrahi, "Multi-objective evolutionary algorithms," encyclopedia of artificial intelligence, (Eds. J. R. Rabuñal, J. Dorado & A. Pazos), Idea Group Publishing, vol. 3, pp. 1145 – 1151,2008.
[15] K. Deb, S. Agrawal., A. Pratap, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE Transactions on Evolutionary Computation, vol. 6, no.2, pp 182-197, 2002.
[16] S. S. Rao, Optimization Theory and Application,New Delhi: Wiley Eastern Limited, 1991.
[17] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989.
[18] Sidhartha Panda "Multi-objective evolutionary algorithm for SSSC-based controller design", Electric Power System Research. vol. 79, issue 6, pp. 937-944, 2009.

Antimicrobial Activity of zero-valent Iron Nanoparticles

Saba A.Mahdy*,Qusay Jaffer Raheed**, P.T. Kalaichelvan
*Centre forAdvancedStudiesinBotany,UniversityofMadras,GuindyCampus,Chennai 600025,TamilNadu,India
**Chemical engineering department, University of Technology, Baghdad, Iraq

ABSTRACT

This work reports on the toxicity of ZVIN nanoparticles on gram-negative and gram-positive bacterial systems, Escherichia coli and Staphylococcus aureus. Detailed characterization of the nanoparticles using x-ray diffraction (XRD), scan electron microscopy (SEM) confirmed the presence of 31.1nm sized ZVIN particles. Further, St. aureus, E.coli were grown in the presence of different ZVIN nanoparticles concentrations for 24 hours. MTT assays were performed and the results provide evidence that ZVIN nanoparticles. FeO nanoparticles MIC of E. coli and St. aureus at concentrations 30 µg/ml, where as growt completely inhibited at concentrations 60 µg/ml.

Keywords- Nanotechnology, ZVIN, bactericide effect, Staphylococcus aureus, Escherichia coli.

I. INTRODUCTION

In the rapidly emerging field of nanopiotechnology, metal nanoparticles are extensively used in drug delivery [1], biosensors [2], bio imaging [3], antimicrobial activities [4], food preservation [5] etc. by exploiting their unique physical chemical and biological properties. There has been a great interest in using microorganisms as a tool for synthesis of new functional inorganic nanomaterials [6, 7] which are free from any kind of toxic chemicals and byproducts. Iron oxide (IO) has been widely used in biomedical research because of its biocompatibility and magnetic properties. [8] IO nanoparticles, with sizes less than 100 nm, have been developed as contrast agents for magnetic resonance imaging (MRI), [9, 10] as hyperthermia agents, [11, 12] and as carriers for targeted drug delivery to treat several types of cancer. [13,14] It is further believed that through the use of magnetic nanoparticles, an optimal drug delivery system can be developed by using an external magnetic field to direct such nanoparticles to desirable sites (such as implant infection) for immediate treatment.

Several recent studies have reported on the antimicrobial activity of nanoparticulate zero-valent iron (ZVIN) [15, 16]. We previously found that ZVIN exhibited a stronger antimicrobial activity than other iron-based nanoparticles, and that the inactivation of E. coli, Staphylococcus aureus by ZVIN was greater under deaerated than air-saturated conditions [17].

The scope of the present study is the synthesis of NZVI particles from ferrous sulfate, and were characterized using scanning electron microgram (SEM) and X-ray diffraction (XRD).and study its antimicrobial activity against St. aureus and E. coli in MTT assay.

II. MATERIAL AND METHODS PREPARATION OF NZVI

The NZVI particles were synthesized by the well-known liquid phase reduction method [18-19]. 10 mmol (2.78 gm) of $FeSO_4 .7H_2O$ was dissolved in 100 ml of an aqueous solution of ammonium persulphate $((NH_4)_2S_2O_8)$. 1.85 g of sodium borohydrate $(NaBH_4)$ was dissolved in 50 ml of distilled water. This solution was added drop wise into the above solution. After addition, this reaction continued for 5 h with constant stirring. The solution was centrifuged for 10 min at 6000 rpm and the supernatant was discarded. The pellet was washed with ethanol three and then dried in vacuum.

MICROORGANISMS AND CULTURE MEDIA

Pure cultures of E. coli (MTCC 118) and St. aureus (MTCC 96) were obtained from Microbial Type Culture Collection (India), were inoculated in 50 mL of Mullar Hinton Broth (Difco Co., Detroit, Mich.) medium and grown at 37°C for 18 h. According to the standard curve correlating bacteria number with optical density, this value was equivalent to 5×10^6 cells/mL

Antimicrobial activity of ZVIN nanoparticles by MTT assay

To measure the activity of living cell by assessing the activity of the bacterial dehydrogenase enzymes. 95µl of the freshly prepared Muller Hinton Broth and the different concentration ofZNVI nanoparticles (10µg, 20µg……100µg) were added and the plates were kept for incubation at 37ºC for 24 hours. 5mg of MTT(3- (4, 5 –Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide) was weighed and dissolved with 1 ml of milli Q water and 10 µl of this preparation is added to each well and kept for 4 hours incubation. The contents were collected and centrifuged at 8000 rpm for 15 minutes, and the pellets were dissolved with 100µl of Dimethyl sulphoxide. Then the contents were transferred to the appropriate well and read at 570 nm in the ELISA reader. The percentage of viable cells was calculated using the following formula.

$$\% \text{ Dead cells} = \frac{\text{Control O.D} - \text{Test O.D}}{\text{Control O.D}} \times 100$$

III- FIGURES AND TABLES
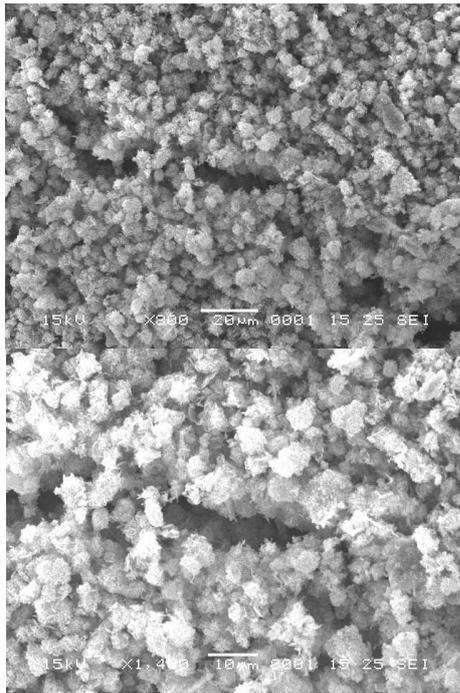
Fig.1 SEM images of NZVa: at 20 µm / 200 X. b: at 10 µm / 1000 X
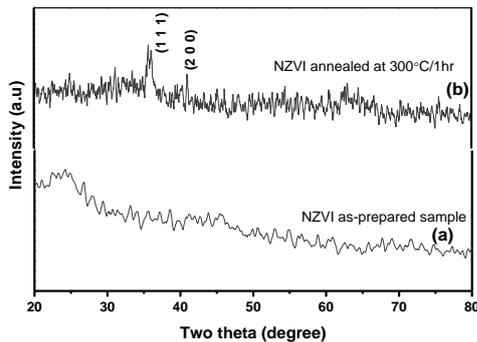

Fig.2 X-Ray Diffraction analysis of NZVI samples.
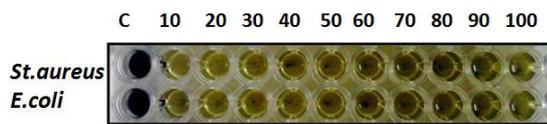

Fig.3 The effect of ZVIN on bacterial growth in MTT plate


.
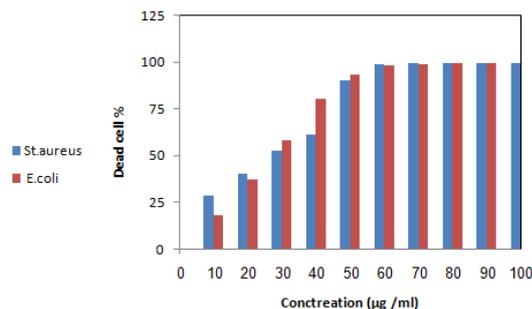
Fig.4 Antimicrobial Activity

IV. RESULTS AND DISCUSSION

NZVI SYNTHESIS AND CHARACTERIZATION

The scanning electron microscopy (SEM) image of synthesized NZVI particles is shown in Fig 2. Results indicate that the synthesized NZVI particles are almost spherical. Fig.2 (a) shows evenly distributed spherical particles approximately 2µm in size, and Fig.2 (b), under higher magnification, confirms the spherical shape and the size range of each particle. On the spherical particles there were threads-like or tube-like structures clearly visible in Fig.2 (b). These structures increased the available surface area of reaction. Fig.3 (a) shows the X-ray diffraction pattern of the as-prepared NZVI sample. The as-prepared samples are amorphous, as no diffraction peaks appear. Fig.3 (b) shows the X-ray diffraction pattern of NZVI annealed at 300°C in air for 1 h. The spectrum shows two major diffraction intensity peaks at $2\theta = 36.08°$ and $41.01°$. The peaks were identified to originate from the (1 1 1) and (2 0 0) planes of FeO respectively (JCPDS no: 772355). The X – ray could be indexed to the $Fm\bar{3}m$ (225) face group (Face – centered) cubic structure, with cell parameter a = 4.309 Å.

The information of the particle size was obtained from the full width at half maximum (FWHM) of the diffracted beam using the sherre: The crystalline size is calculated using the Debye-Scherrer formula :

$$D = \frac{0.9\,\lambda}{\beta \cos\theta}$$

The sample annealed at 300°C/1h has an average crystalline size of 31.1nm ± 0.5. As the annealing time or temperature increases, the crystalline size increases.

The antimicrobial activity of FeO nanoparticles on bacterial growth are shown in figure 3 was determined by MTT assay. In this experiment E.coli and St.aureus test strains were inoculated in Muller Hinton medium supplemented with different concentrations of FeO nanoparticles. The MIC values for the FeO nanoparticles were 30µg/ml, for both. Increasing concentration of FeO nanoparticles substantially inhibited the growth of E.coli and St.aureus test strains completely inhibited in 60 µg/ml shown in figure 4.

There are several factors that caused the presently studied FeO nanoparticles to be bactericidal. The main mechanism by which antibacterial drugs and antibiotics work is via oxidative stress generated by ROS.[20]ROS, including superoxide radicals(O_2-), hydroxyl radicals (-OH), hydrogen peroxide (H_2O_2), and singlet oxygen ($1O_2$), can cause damage to proteins and DNA in bacteria.[21] In this case, metal oxide FeO could be the source that created ROS leading to the inhibition of St. aureus. A similar process was described by Keenan et al in which Fe^{2+} reacted with oxygen to create hydrogen peroxide. This H_2O_2 consequently reacted with ferrous irons via the Fenton reaction and produced hydroxyl radicals which are known to damage biological macromolecules.[22] Other research has demonstrated

that the small size of nanoparticles can also contribute to bactericidal effects. For example, Lee et al reported that the inactivation of Escherichia coli by zero-valent iron nanoparticles [23] could be because of the penetration of the small particles (sizes ranging from 10–80 nm) into E. coli membranes. Nano-Fe0 could then react with intracellular oxygen, leading to oxidative stress and eventually causing disruption of the cell membrane. Several other studies on ZnO and MgO nanoparticles also concluded that antibacterial activity increased with decreasing particle size. [24,25] In this study, the concentration of nanoparticles was a major contribution to St. aureus activity inhibition. A similar concentration-dependent behavior was observed by Kim et al when they investigated the antimicrobial effects of Ag and ZnO nanoparticles on St. aureus and E. coli. [24,25] In a study of bactericidal effects of IO nanoparticles on St. epidermidis, Taylor et al also reported concentration dependent bacteria inhibition.[26] Briefly, St. epidermidis density progressively decreased at time points of 12, 24, and 48 hours when incubated with 100 μg/mL, 1 mg/mL, and 2 mg/mL IO. It is also important to note that IO nanoparticles do not negatively influence all cells. Specifically, osteoblast (boneforming cell) proliferation was enhanced in the presence of Fe2O3 nanoparticles (at 4.25 mg/mL. [27] Such results showed that FeO nanoparticles could have a dual therapeutic function which can enhance bone growth and inhibit bacterial infection. Lastly, this present study provided evidence that with an appropriate external magnetic field, FeO nanoparticles may be directed to kill bacteria as needed throughout the body.

V. CONCLUSION

Stable FeO nanoparticles were successfully synthesized. The particles were characterized with SEM, dynamic light scattering, XRD. A live/dead assay showed that at the highest dose of iron oxide (30μg/mL), the growth of E. coli and S. aureus was inhibited significantly compared with the control samples. Indicates that ZVIN have potential for use as antimicrobials. ZVIN has several advantages such as low cost, easy preparation, and high reactivity compared to other metal nanoparticles.

REFERENCES

[1] Gupta AK, Gupta M. Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. Biomaterials. 2005; 26(18):3995–4021.

[2] Berry CC, Curtis ASG. Functionalisation of magnetic nanoparticles for applications in biomedicine. J Phys D Appl Phys. 2003;36: R198–R206.

[3] Beets-Tan RGH, Van Engelshoven JMA, Greve JWM. Hepatic adenoma and focal nodular hyperplasia: MR f indings with superparamagnetic iron oxide- enhanced MRI. Clin Imaging. 1998;22(3): 211–215.

[4] Babes L, Denizot B, Tanguy G, Le Jeune JJ, Jallet P. Synthesis of iron oxide nanoparticles used as MRI contrast agents: A parametric study J Colloid Interface Sci. 1999;212(2):474–482.

[5] Chan DCF, Kirpotin DB, Bunn PA. Synthesis and evaluation of colloidal magnetic iron oxides for the site-specific radio frequency induced hyperthermia of cancer. J Magn Mag Mat. 1993;122(1–3): 374–378.

[6] ALT V, BECHERT T, STEINRCKE P, WAGENER M, SEIDEL P, DINGELDEIN E, DOMANN U, SCHNETTLER R. AN IN vitro assessment of the antibacterial properties and cytotoxicity of nanoparticulate silver bone cement. Biomaterials 2004; 25:4383–4391. [PubMed: 15046929].

[7] Stoimenov PK, Klinger RL, Marchin GL, Klabunde KJ. Metal oxide nanoparticles as bactericidal agents. Langmuir 2002;18:6679–6686.

[8] Matheson LJ, Tratnyek PG. Reductive dehalogenation of chlorinated methanes by iron metal. Environ Sci Technol 1994; 28:2045–2053.

[9] Farrell J, Kason M, Melitas N, Li T. Investigation of the long-term performance of zero-valent iron for reductive dechlorination of trichloroethylene. Environ Sci Technol 2000;34:514–521.

[10] Joo SH, Feitz AJ, Sedlak DL, Waite TD. Quantification of the oxidizing capacity of nanoparticulate zero-valent iron. Environ Sci Technol 2005;39:1263–1268. [PubMed: 15787365].

[11] Zhang WX. Nanoscale iron particles for environmental remediation: An overview. J Nanopart Res 2003;5:323–332.

[12] Keenan CR, Sedlak DL. Factors affecting the yield of oxidants from the reaction of nanoparticulate zero-valent iron and oxygen. Environ Technol. 2008;42(4):1262–1267.

[13] Lee C, Kim JY, Lee WI, Nelson KL, Yoon J, Sedlak DL. Bactericidal effect of zero-valent iron nanoparticles on Escherichia coli. Environ Technol. 2008;42(13):4927–4933.

[14] Phenrat, T., T. C. Long, G. V. Lowry, and B. Veronesi.Partial oxidation ("aging") and surface modification decrease the toxicity of nanosized zerovalent iron. Environ. Sci. Technol. 2009; 43:195–200.

[15] Lee, C., and D. L. Sedlak. Enhanced formation of oxidants from bimetallic nickel-iron nanoparticles in the presence of oxygen. Environ. Sci. Technol.2008; 42:8528–8533.

[16] Auffan, M., W. Achouak, J. Rose, M. Roncato, C. Chanéac, D. T. Waite, A. Masion, J. C. Woicik, M. R. Wiesner, and J. Bottero. Relation between the redox state of iron-based nanoparticles and their cytotoxicity toward Escherichia coli. Environ. Sci. Technol.2008; 42:6730– 6735.

[17] Boxall, A.B, Tiede K., Chaudhry Q. Engineered nanomaterials in soils and water: how do they behave and could they pose a risk to human health? Nanomed 2007; 2:919–27.

[18] Christian, P., Von der Kammer, F., Baalousha, M., Hofmann T. Nanoparticles: structure, properties, preparation and behaviour in environmental media. Environ Toxicol Chem. 2008; 17:326–43.

[19] Liang,J.F., Guo,Y., Fan,M., Wang,J., Yang, H. Reduction of Nitrite by Ultrasound-Dispersed Nanoscale Zero-Valent Iron Particles. *Ind. Eng. Chem. Res. 47,* 2008*;* 8550–8554.

[20] Sies H. Oxidative stress: oxidants and antioxidants. Exp Physiol.1997;82(2):291–295.

[29] Park HJ, Kim JY, Kim J, et al. Silver-ion-mediated reactive oxygen species generation affecting bactericidal activity. Water Res. 2009;43(4):1027–1032.

[21] Kim JS, Kuk E, Yu KN, et al. Antimicrobial effects of silver nanoparticles. Nanomed Nanotechnol Biol Med. 2007;3(1):95–101.

[22] Touati D. Iron and oxidative stress in acteria. Arch Biochem Biophys.2000;373(1):1 6.

[23] Lee C, Kim JY, Lee WI, Nelson KL, Yoon J, Sedlak DL. Bactericidal effect of zero-valent iron nanoparticles on Escherichia coli. Environ Technol. 2008;42(13):4927–4933.

[24] Zhang L, Jiang Y, Ding Y, Povey M, York D. Investigation into the antibacterial behaviour of suspensions of ZnO nanoparticles (ZnO nanofluids). J Nanopart Res. 2007;9(3):479–489.

[25] Makhluf S, Dror R, Nitzan Y, Abramovich Y, Jelinek R, Gedanken A. Microwave-assisted synthesis of nanocrystalline MgO and its use as a bacteriocide. Adv Funct Mater. 2005;15(10):1708–1715.

[26] Taylor EN, Webster TJ. The use of superparamagnetic nanoparticles for prosthetic biofilm prevention. Int J Nanomedicine. 2009;4:145–152.

[27] Pareta RA, Taylor E, Webster TJ. Increased osteoblast density in the presence of novel calcium phosphate coated magnetic nanoparticles.Nanotechnology.2008;19(26):265101.